

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 20:22:40 ; Search time 2633 Seconds

(without alignments)
12498.758 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 1 ggcctcagagccagatgcct.....ataagaaaaaaaaaaaaa 2032

Sequence: 1 ggcctcagagccagatgcct.....ataagaaaaaaaaaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 25

Total number of hits satisfying chosen parameters: 24
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	31.0	954	BI823850	BI823850 603039012
2	581	28.6	593	BM969282	BM969282 DT-CF-ENO
3	551	27.1	571	BM129080	BM129080 I117C04.Y
4	478	23.5	553	BM128370	BM128370 I113d12.X
5	478	23.5	553	BM128831	BM128831 I117C04.X
6	468	23.0	668	AL709927	AL709927 DKPZ6860

7	318	15.6	680	12	BF878439	BF878439 MRO-ET010
C	288	14.2	362	10	AM002418	AM002418 wu61g03.X
8	283	13.9	419	10	AM572390	AM572390 xu02a06.X
C	150	7.4	417	10	AM572510	AM572510 xq16g11.X
C	136	6.7	358	12	BF056840	BF056840 7K10d12.X
C	112	5.9	525	12	BF197521	BF197521 7084a08.X
C	59	2.9	620	9	A1824100	A1824100 w146c01.X
C	59	2.9	695	12	BE857538	BE857538 7g01a08.X
C	15	2.3	735	12	BE858652	BE858652 7g01a09.X
C	16	2.3	436	10	AK081348	AK081348 xc41b06.X
17	27	1.3	497	13	BM246681	BM246681 K0741E04.X
18	27	1.3	536	13	BM245312	BM245312 K0722H07.X
19	27	1.3	695	13	BG964671	BG964671 602831875
20	27	1.3	783	13	BG963298	BG963298 602827716
21	27	1.3	852	13	BG966340	BG966340 602832826
22	27	1.3	1923	11	AK009113	AK009113 MUS muscu
23	25	1.2	501	13	BM036473	BM036473 Iub107.X
C	24	1.2	682	17	CNS028KQ	AL189971 Tetradon

ALIGNMENTS

RESULT 1
BI823850 954 bp mRNA linear EST 04-OCT-2001
DEFINITION 603039012F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179826 5',
LOCUS BI823850
ACCESSION BI823850
VERSION BI823850.1 GI:15935400
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 954)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: L1AM11448 row: h column: 03
High quality sequence stop: 856.
Location/Qualifiers

FEATURES

1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179826"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pcwv-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT

199 a 297 c 250 g 208 t

Query Match 31.0%; Score 630; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 1,1e+186;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 160 CAAGGCTTCACCTGAGCACAATGCTAGCTGCTAAATAAATGAAAGCTGCTGTTTCT 219
    |||
    |||
    |||
Db 56 CAAGGCTTCACCTGAGCACAATGCTAGCTGCTAAATAAATGAAAGCTGCTGTTTCT 115
    |||
    |||
    |||
OY 220 GGTTCACCAATGGCCATCTTGGCTCTATTCTTCACATGTACAGCCACACATCAGCTC 279
    |||
    |||
    |||
Db 116 GGTTCACCAATGGCCATCTTGGCTCTATTCTTCACATGTACAGCCACACATCAGCTC 175
    |||
    |||
    |||
OY 280 CCGTGTATGAAGGACAGCCGAGGCGCATGACGAGCGGTCTGCTTCCGGGCGCTC 339
    |||
    |||
    |||
Db 176 CCGTGTATGAAGGACAGCCGAGGCGCATGACGAGCGGTCTGCTTCCGGGCGCTC 235
    |||
    |||
    |||
OY 340 TGGCTCTCTTTTGTGGGGCAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGA 399
    |||
    |||
    |||
Db 236 TGGCTCTCTTTTGTGGGGCAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGA 295
    |||
    |||
    |||
OY 400 GCGCGCTGGCAGCTGTGATGTACCTTCACAGCAGACCGGCTGATGCTGACATGGC 459
    |||
    |||
    |||
Db 296 GCGCGCTGGCAGCTGTGATGTACCTTCACAGCAGACCGGCTGATGCTGACATGGC 355
    |||
    |||
    |||
OY 460 TGTGGGATCTGATAGCGGCGCTCTTCTTGTGCGACATGAGGCTTTCATGCTACAT 519
    |||
    |||
    |||
Db 356 TGTGGGATCTGATAGCGGCGCTCTTCTTGTGCGACATGAGGCTTTCATGCTACAT 415
    |||
    |||
    |||
OY 520 GGAACCTGGTCCCGGAGACAGTCACAGCTCTTTCAGTGGGAGAACAGCCGGGCGCTGTG 579
    |||
    |||
    |||
Db 416 GGAACCTGGTCCCGGAGACAGTCACAGCTCTTTCAGTGGGAGAACAGCCGGGCGCTGTG 475
    |||
    |||
    |||
OY 580 TTTCGACCTGCTGTGATCATCCACAGATGAATATGATCCCGGGGCTACTGACG 639
    |||
    |||
    |||
Db 476 TTTCGACCTGCTGTGATCATCCACAGATGAATATGATCCCGGGGCTACTGACG 535
    |||
    |||
    |||
OY 640 GCTTCGTGAGTACACAGCCCTTGAAGTGTGTGAGAGAGCGCTGCTCTACAGCCA 699
    |||
    |||
    |||
Db 536 GCTTCGTGAGTACACAGCCCTTGAAGTGTGTGAGAGAGCGCTGCTCTACAGCCA 595
    |||
    |||
    |||
OY 700 CGTGTGCTCAGAGAGGTGGCTTCTTCACAGCTGACAGTCCCTTACCGGCTGTGAAAG 759
    |||
    |||
    |||
Db 596 CGTGTGCTCAGAGAGGTGGCTTCTTCACAGCTGACAGTCCCTTACCGGCTGTGAAAG 655
    |||
    |||
    |||
OY 760 CCCCTCCCTCAACCTGCATATCGTGACCT 789
    |||
    |||
    |||
Db 656 CCCCTCCCTCAACCTGCATATCGTGACCT 685
    |||
    |||
    |||

RESULT 2
BM969292/c 593 bp mRNA linear EST 21-MAR-2002
LOCUS BM969292
DEFINITION UI-CF-ENO-acp-1-21-0-UI-s1 UI-CF-ENO Homo sapiens cDNA clone
VERSION BM969292
ACCESSION BM969292.1 GI:19586879
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS Ronaldo M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
        discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: McCray, PB
        University of Iowa
        2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
        Tel: 319 356 4866
        Fax: 319 356 7171
        Email: paul-mccray@uiowa.edu
        Tissue Procurement: Dr. M. J. Welsh, University of Iowa
        cDNA library preparation: Dr. M. Bento Soares, University of Iowa

```

FEATURES

Source

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Cloning Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA-Yes.

Location/Qualifiers

```

1..593
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-ENO-acp-1-21-0-UI"
/clone_lib="UI-CF-ENO"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
UI-CF-ENO is a cDNA library containing the following
tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells.
The library was constructed according to Ronaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dr)18 tail. The sequence tag for this
library is CTCCTCAGGT.
TAG_Lib=UI-CF-ENO
TAG_Tissue=Human Lung Epithelial cell lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTCCTCAGGT"
BASE COUNT 157 a 127 c 137 g 172 t
ORIGIN

```

Query Match 28.6%; Score 581; DB 14; Length 593;
 Best Local Similarity 100.0%; Pred. No. 2,7e-171;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1436 ACATGCTGTGGGATACACTGATGTGAGTGTGTCACACGTCGTCACAGCAGAGA 1495
    |||
    |||
    |||
Db 593 ACATGCTGTGGGATACACTGATGTGAGTGTGTCACACGTCGTCACAGCAGAGA 534
    |||
    |||
    |||
OY 1496 CTTTGTGTCATGCTGTGCTAGAAAACAGACGTGGGAACTATGTAGCAGACAT 1555
    |||
    |||
    |||
Db 533 CTTTGTGTCATGCTGTGCTAGAAAACAGACGTGGGAACTATGTAGCAGACAT 474
    |||
    |||
    |||
OY 1556 CCCACAGTGAACAGAGGTATTGCTCTTCTTTCTTGTGATCTTCTGTCGGCAGAC 1615
    |||
    |||
    |||
Db 473 CCCACAGTGAACAGAGGTATTGCTCTTCTTTCTTGTGATCTTCTGTCGGCAGAC 414
    |||
    |||
    |||
OY 1616 TTCAGAGCTTTGSGCCGCGAGGCGCTATTAGCAGCAGACAGTACAGTGAATGATC 1675
    |||
    |||
    |||
Db 413 TTCAGAGCTTTGSGCCGCGAGGCGCTATTAGCAGCAGACAGTACAGTGAATGATC 354
    |||
    |||
    |||
OY 1676 CATAAACCTCCCTGCCATCTTCCCATGAGGGAATGATCTTTCACAAAGAGCTCA 1735
    |||
    |||
    |||
Db 353 CATAAACCTCCCTGCCATCTTCCCATGAGGGAATGATCTTTCACAAAGAGCTCA 294
    |||
    |||
    |||
OY 1736 CCAGCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGGAGTTCCTCATGATTAAGA 1795
    |||
    |||
    |||
Db 293 CCAGCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGGAGTTCCTCATGATTAAGA 234
    |||
    |||
    |||
OY 1796 AGGAAGTGGAGAACAGGTGTGAGTCTTATGAGCTTGAACATCAGATATGCTTA 1855
    |||
    |||
    |||
Db 233 AGGAAGTGGAGAACAGGTGTGAGTCTTATGAGCTTGAACATCAGATATGCTTA 174
    |||
    |||
    |||
OY 1856 TCAGAAATATGAACAAAATCTCTGACAAAGAGCAGAGCTTTAAGTTACAGGGTGCC 1915
    |||
    |||
    |||
Db 173 TCAGAAATATGAACAAAATCTCTGACAAAGAGCAGAGCTTTAAGTTACAGGGTGCC 114
    |||
    |||
    |||

```

QY 1916 TGGGCTGCATTGTAATATACCTCCCTCTGCATTTTCCCATACATAGAGACTTTGAC 1975
 |||||||
 Db 113 TGGGCTGCATTGTAATATACCTCCCTCTGCATTTTCCCATACATAGAGACTTTGAC 54
 |||||||
 QY 1976 CTGTGAAGCTGCATCTGTTAATAGTAATATCCCAATAA 2016
 |||||||
 Db 53 CTGTGAAGCTGCATCTGTTAATAGTAATATCCCAATAA 13
 |||||||

RESULT 3

BM129080 571 bp mRNA linear EST 12-MAR-2002
 LOCUS 117C04.Y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 DEFINITION cdna clone IMAGE:5676655 5' similar to TR:09Y5R3 Q9Y5R3
 N-ACETYLGLUCOSAMINE 6-O-SULFOTRANSFERASE ; mRNA sequence.

ACCESSION BM129080 GI:17123632
 VERSION BM129080
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 511)
 AUTHORS Mammalia, Euthera; Primates; Catarrhini; Homnidae; Homo.
 Lemisha, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,
 Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: 117C04.x1

TITLE Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 JOURNAL Endocrine Pancreas Consortium
 COMMENT Other ESTs: 117C04.x1

DEPT OF Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For Information on
 obtaining a clone please contact: Juliana Brown
 (brownj@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@image.llnl.gov
 High quality sequence stop: 434.

FEATURES
 SOURCE location/Qualifiers
 1..571
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5676655"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /note="Organ: Pancreas; Vector: pSPORT1, Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). CDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an EcoT of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

BASE COUNT 109 a 169 c 143 g 150 t

ORIGIN

Query Match 27.1%; Score 551; DB 13; Length 571;
 Best Local Similarity 100.0%; Pred. No. 6, 1e-162;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 TGAGTCTAGAGAGTGTAAAGCTGTGTTACTTTCACAGCTTCCTGGAGAGAGTGTCTCTC 147
 Db 1 TGAGTCTAGAGAGTGTAAAGCTGTGTTACTTTCACAGCTTCCTGGAGAGAGTGTCTCTC 60
 QY 148 AAGCCCGCTTTCAGAGTGTTCACATTCAGACATGCTACTGCTAAATAATGAAGCT 207
 Db 61 AAGCCCGCTTTCAGAGTGTTCACATTCAGACATGCTACTGCTAAATAATGAAGCT 120
 QY 208 CCGTGTCTTGTGTTTCCAGATGGCAGTCTGCTATCTTCCAGATGTCAGAGCA 267
 Db 121 CCGTGTCTTGTGTTTCCAGATGGCAGTCTGCTATCTTCCAGATGTCAGAGCA 180
 QY 268 CAACATCAGCTCCCTGTCTATGAAGGACACAGCCGAGGAGATGCACTGCTGTCTCTC 327
 Db 181 CAACATCAGCTCCCTGTCTATGAAGGACACAGCCGAGGAGATGCACTGCTGTCTCTC 240
 QY 328 TTCTCGGCGCTCTGCTCTCTCTTTTGTGGGAGCTTTTGGGAGACCCAGATGTTT 387
 Db 241 TTCTCGGCGCTCTGCTCTCTCTTTTGTGGGAGCTTTTGGGAGACCCAGATGTTT 300
 QY 388 CTACCTGATGAGCCCGGCTGAGAGTGTGATGACCTTCAAGAGAGACCGCTGGAT 447
 Db 301 CTACCTGATGAGCCCGGCTGAGAGTGTGATGACCTTCAAGAGAGACCGCTGGAT 360
 QY 448 GCTGCATAGTCTGTGGGAGATCTGATACGAGCGCTCTTGTGCGACATGAGCGTCT 507
 Db 361 GCTGCATAGTCTGTGGGAGATCTGATACGAGCGCTCTTGTGCGACATGAGCGTCT 420
 QY 508 TGATGCTACATGAGACCTGTGTCGCGAGAGAGCCGCTTTCAGTGGAGAGAG 567
 Db 421 TGATGCTACATGAGACCTGTGTCGCGAGAGAGCCGCTTTCAGTGGAGAGAG 480
 QY 568 CCGGCGCTGTGTTCTGACCTGCTGTGACATCATCCACAGATGATATCATCCCCG 627
 Db 481 CCGGCGCTGTGTTCTGACCTGCTGTGACATCATCCACAGATGATATCATCCCCG 540
 QY 628 GGCTCAGTCA 638
 Db 541 GGCTCAGTCA 551

RESULT 4
 LOCUS BM128370/c 553 bp mRNA linear EST 12-MAR-2002
 DEFINITION 1133d1.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 ACCESSION BM128370
 VERSION BM128370.1 GI:17122922
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 553)

REFERENCE Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemisha, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,
 Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)

TITLE Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 JOURNAL Endocrine Pancreas Consortium
 COMMENT Other ESTs: 1133d1.x1

DEPT OF Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557

QY	1540	TATGTGAGCAGCAATCCCAACAGTAACAGGGATTTGCTCTCTCTCTCTCTGACT	1599
Db	553	TATGTGAGCAGCAATCCCAACAGTAACAGGGATTTGCTCTCTCTCTCTCTGACT	494
QY	1600	TCTGTCTGGGCACTTCCAGACCTTTGTGCTGTGGAGGCTATTAAACACACAGT	1658
Db	493	TCTGTCTGGGCACTTCCAGACCTTTGTGCTGTGGAGGCTATTAAAGCACACAGT	434
QY	1660	ATGAGTGGATTATGCCATTAACCTCCCTGTCCACATCTTGCGCAATGGGGAATGACAT	1719
Db	433	ATGAGTGGATTATGCCATTAACCTCCCTGTCCACATCTTGCGCAATGGGGAATGACAT	374
QY	1720	TTCCACCAAGAGCTCACACAGATTTTCCACAGAGATGCAAATTCGAGCCCTTGAGTTC	1779
Db	373	TTCCACCAAGAGCTCACACAGATTTTCCACAGAGATGCAAATTCGAGCCCTTGAGTTC	314
QY	1780	CCAGTGATTCGAAGGAAGAAGTGGGAACAAGTGTGATGCTACTTATGACTTGACCA	1838
Db	313	CCAGTGATTCGAAGGAAGAAGTGGGAACAAGTGTGATGCTACTTATGACTTGACCA	254
QY	1840	TCACAGCTATCGGTATCATCAAGAAATATGAACAAATAATCTGTGCACAAAAGACAAGCTTT	1899
Db	253	TCACAGCTATCGGTATCATCAAGAAATATGAACAAATAATCTGTGCACAAAAGACAAGCTTT	194
QY	1900	AAGTTCACAGGGTGCCTGGGCTGCATTTTGAATATCACTTCCCTCTGCATTTCCACATCA	1959
Db	193	AAGTTCACAGGGTGCCTGGGCTGCATTTTGAATATCACTTCCCTCTGCATTTCCACATCA	134
QY	1960	CATAGGAAGCTTTGACCTGTGANAAGCTGCATCTGTTAAATACATAAATTTCCAAAATAG	2017
Db	133	CATAGGAAGCTTTGACCTGTGANAAGCTGCATCTGTTAAATACATAAATTTCCAAAATAG	76

RESULT 5
BM128831/c

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:567655"
/clone_id="Melton Normalized Human Islet 4 N4-HIS 1"
/sec="both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPOR1; Site:1: Not 1; Site:2: Sal 1; Starting library constructed using SuperScript plasmid library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 genome Research 6:791-806. 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

BASE COUNT	163 a	111 c	127 g	152 t
ORIGIN				
Query Match		23.5%	Score 478:	DB 13: Length 553:
Best Local Similarity		100.0%	Pred. No. 3.3e-139;	
Matches 478; Conservative		0;	Mismatches 0;	Indels 0; Gaps
QY 1540	TATGTGAGCAGACATATCCACACAGTGAACAGGGATTTGCTCTCTCTTTCTTGATCT			1599
Db 553	TATGTGAGCAGACATATCCACACAGTGAACAGGGATTTGCTCTCTCTCTTTCTTGATCT			494
QY 1600	TCCGCTGTGGGCGACACTTCGAGACATTTGTGGCTGTGGAGCCCTATTAAACAGACACAGT			1655
Db 493	TCCGCTGTGGGCGACACTTCGAGACATTTGTGGCTGTGGAGCCCTATTAAACAGACACAGT			434
QY 1660	ATCAGTGAATTATCCATAAACCTCCCTGTCCACATCTTGTCCACATGGGGAATGATCT			1719

Db	Accession	Version	Keywords	Source	Organism	Reference	Authors	Features	Source
Db	433	ATCAGTGAATTTGATCCATAAACCTCCCTGTCACATCTTGCCCAATGGGAATGATTC							
QY	1720	TTCCACCAAGAGCTTCACAGATTTTCCACAGAGATGCAATTTGACCCCTTGAGTTC							
Db	373	TTCCACCAAGAGCTTCACAGATTTTCCACAGAGATGCAATTTGACCCCTTGAGTTC							
QY	1780	CCAGTGGATTCAGAGAGAAAGTGGGAACAGGTGTGATGCTTACTTATGAGCTTGACCA							
Db	313	CCAGTGGATTCAGAGAGAAAGTGGGAACAGGTGTGATGCTTACTTATGAGCTTGACCA							
QY	1840	TCACAGCATATGCGTATATGASAAATATGASAAATATCTCTGACAAAGAGCAACTCTT							
Db	253	TCACAGCATATGCGTATATGASAAATATGASAAATATCTCTGACAAAGAGCAACTCTT							
QY	1900	AAGTTCACAGAGGTGCGCTGGGCTGCATTTGAAATATACATTCGCCCTGCAATTTCCATCA							
Db	193	AAGTTCACAGAGGTGCGCTGGGCTGCATTTGAAATATACATTCGCCCTGCAATTTCCATCA							
QY	1960	CATGAAACACTTTGACCGCTGTAAGCGTGCACATCTGTATATCTAAATCCCAATAAG							
Db	133	CATGAAACACTTTGACCGCTGTAAGCGTGCACATCTGTATATCTAAATCCCAATAAG							
RESULT 6	AL709927	668 bp	mRNA	linear	EST 22-MAR-2002				
LOCUS	DKFZP68602364.f1	686	(synonym: hlcc3)	Homo sapiens	cdna clone				
DEFINITION	DKFZP68602364.5		mRNA sequence.						
ACCESSION	AL709927								
VERSION	AL709927.1	GI:19693282							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;								
AUTHORS	Mammalia: Euthera: Primates; Catarrhini; Homnidae; Homo.								
	1 (bases 1 to 668)								
	Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann								
	J., S.								
	EST (Duesterhoeft, et al.)								
	Unpublished (1999)								
	Contact: Duesterhoeft A								
	MPs								
	Am Kioferpsitz 18a D-82152 Martinsried, Germany								
	This is the 5' sequence of the clone insert								
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer								
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;								
	sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing								
	consortium of the German Genome Project.								
	No sl sequence available.								
	This clone (DKFZP68602364) is available at the RZPD in Berlin.								
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059								
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.								
	Location/Qualifiers								
	1..668								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/clone="DKFZP68602364"								
	/clone_1ib="686 (synonym: hlcc3)"								
	/tissue_type="human skeletal muscle"								
	/dex_strage="adult"								
	/lab_host="DH10B"								
	/note="Vector: pT7-1Ex2; Site_1: SfilI; Site_2: SfilI;								
	CDNA-collection"								
	BASE COUNT	132 a	202 c	178 g	154 t	2 others			
	ORIGIN								
	Query Match	23.0%	Score 468	DB 9	Length 668				

[illegible]

```

RESULT 8
AM002418/c
LOCUS
DEFINITION
AM002418
Accession
AM002418
VERSION
AM002418.1
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 362)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/BLAT at:
www-bio.lnl.gov/dbfp/image/image.html
Seq primer: 40UP from GIBCO.
Location/Qualifiers
1..362
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```

QY	1730	AGCTACACAGCATTTTCCACAGCAATGCMAATTCGTGACCCCTGGAGTCCCAATGATTT	1789
Db	362	AGCTACACAGCATTTTCCACAGCAATGCMAATTCGTGACCCCTGGAGTCCCAATGATTT	303
QY	1790	CAAGGAAGGAATGGGAACAAGTTGGATGCTACTTATGAGCTTGGACCATCAGCCTAT	1849
Db	302	CAAGGAAGGAATGGGAACAAGTTGGATGCTACTTATGAGCTTGGACCATCAGCCTAT	243
QY	1850	CGGTATACAGAAATATGAAACAAATCTCTCACAAAAGACAGACGCTTAAGTTCCAG	1909
Db	242	CGGTATACAGAAATATGAAACAAATCTCTCTCACAAAAGACAGACGCTTAAGTTCCAG	183
QY	1910	GGTGCCTGGGCTGCATTTGAAATACACTTCCCTCTGCTATTTCCATCACAATGAAAGAC	1969
Db	182	GGTGCCTGGGCTGCATTTGAAATACACTTCCCTCTGCTATTTCCATCACAATGAAAGAC	123
QY	1970	TTTACCGTGAAGCTGCATCTCTTAACTATAAATTTCCCAATTAAG	2017
Db	122	TTTACCGTGAAGCTGCATCTCTTAACTATAAATTTCCCAATTAAG	75

```

RESULT 9
LOCUS      AM572390/c
DEFINITION AM572390 419 bp mRNA linear EST 13-MAR-2000
            XM09a06.x2 NCI-CGAP_Col4 Homo sapiens CDNA clone IMAGE:2795634 3',
            mRNA sequence.
ACCESSION  AM572390
VERSION    AM572390.1 GI:7237123
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 419)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaps@remail.nih.gov
            Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.
            Emmer-Buck, M.D., Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/HLNT at:
            image.llnl.gov/image/html/irsources.shtml
            Seq primer: -40UP from Gibco
            High quality sequence stop: 343.
            Location/Qualifiers
                1..419
                /organism="Homo sapiens"

```

BASE COUNT	122 a	78 c	89 g	130 t	
ORIGIN					
Query Match	13.9%	Score 283;	DB 10;	Length 419;	
Best Local Similarity	99.7%	Pred. No. 2e-78;			
Matches 333;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
/db_xref="taxon:9606" /clone="IMAGE:2799634" /clone_id="NCI_CGAP_C014" /tissue_type="moderately-differentiated adenocarcinoma" /lab_host="DH10B" /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: SalI, Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.7 kb. Life Technologies catalog #: 11531-019"					
OY	1684	TCCCGTCCACATCTTGGCCCATGGGGAAATGATCTTTCCACCAAGAGCTCACCGCATTT	1743		
DB	419	TCCCGTCCACATCTTGGCCCATGGGGAAATGATCTTTCCACCAAGAGCTCACCGCATTT	360		
OY	1744	TTCACAGAGATGCAAAATTCATGAGCCCTTGAGGTCCAGATTCAGATTCAGAGAGAGTG	1803		
DB	359	TTCACAGAGATGCAAAATTTTGAAGCCCTTGAGGTCCAGATTCAGAGAGAGTG	300		
OY	1804	GGAACAAGGTGGATGCTTACTTATGAGCTTGACCATACACAGCTATCGGTAATCAGAAAT	1863		
DB	299	GGAACAAGGTGGATGCTTACTTATGAGCTTGACCATACACAGCTATCGGTAATCAGAAAT	240		
OY	1864	ATGAACAACAAATCTCTGACACAAAGAGAGCTTAAAGTTCACAGGGGCTCGGCTGCG	1923		
DB	239	ATGAACAACAAATCTCTGACACAAAGAGAGCTTAAAGTTCACAGGGGCTCGGCTGCG	180		
OY	1924	ATTGTAATATCATCTTCCCTCTGATTTTCCCATCACAATGAGAAGCTTGACCTTGGAAG	1983		
DB	179	ATTGTAATATCATCTTCCCTCTGATTTTCCCATCACAATGAGAAGCTTGACCTTGGAAG	120		
OY	1984	CTGCCATCTGTTAATCTAAATTCGCCCAATAG	2017		
DB	119	CTGCCATCTGTTAATCTAAATTCGCCCAATAG	86		
RESULT 10					
AM572510/c	417 bp	MRNA	linear	EST 13-MAR-2000	
LOCUS	AM572510				
DEFINITION	xq18g11.x2 NCI_CGAP-UCL Homo sapiens cDNA clone IMAGE:2751044 3'				
ACCESSION	AM572510				
VERSION	AM572510.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 417)				
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
	Unpublished (1997)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapds-rt@mail.nih.gov				
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.				
	Emmert-Buck, M.D., Ph.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	DNA Library Arrayed by: Greg Lennon, Ph.D.				
	DNA Sequencing by: Washington University Genome Sequencing Center				
	clone distribution: NCI-CGAP clone distribution information can be				
	found through the I.M.A.G.E. Consortium/INM at:				
	image.lnl.gov/image/html/lifresources.shtml				
	Seq primer: -40UP from G16C				
	High quality sequence stop: 416.				
	Location/Qualifiers				
	1. 417				
FEATURES					
SOURCE	/organism="Homo sapiens"				

BASE COUNT	140 a	65 c	80 g	132 t
ORIGIN				
Query Match	7.4%;	Score 150;	DB 10;	Length 417;
Best Local Similarity	100.0%;	Pred. No. 5,1e-37;		
Matches 150;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1868	AAACAATCTCTGCACAAAAAGCAGCTCTTAAGTTCACACAGGCGTGGCGTCATTT	1927		
DB 383	AAACAATCTCTGCACAAAAAGCAGCTCTTAAGTTCACACAGGCGTGGCGTCATTT	324		
QY 1928	GAATATCACTTCCTCCCTGCATTTCCCATCATCAGAAAGACTTGAGCTGGAAGCTGC	1987		
DB 323	GAATATCACTTCCTCCCTGCATTTCCCATCATCAGAAAGACTTGAGCTGGAAGCTGC	264		
QY 1988	CATCTGTTAATCTAAATTCACAAATAG	2017		
DB 263	CATCTGTTAATCTAAATTCACAAATAG	234		
RESULT 11				
BF056840/c				
LOCUS	BF056840	358 bp	mRNA	linear
DEFINITION	7k10d12.x1 NCI-CGAP-GC6 Homo sapiens CDNA clone IMAGE:3443783 3',			
ACCESSION	BF056840			
VERSION	BF056840.1	GI:10810736		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases:1 to 358)			
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: Christopher A. Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonafio, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -400P from Glbco. Location/Qualifiers 1. .358 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3443783" /clone_lib="NCI-CGAP-GC6" /tissue_type="pooled germ cell tumors" /lab_host="DH10B" /note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP-GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool			

of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT
ORIGIN

107 a 68 c 82 g 101 t

Query Match

6.7%; Score 136; DB 12; Length 358;

Best Local Similarity 100.0%; Pred. No. 1.3e-32; Mismatches 0; Indels 0; Gaps 0;

Matches 136; Conservative 0;

QY 1730 AGCTCACACACATTTTCCACAGAGATGCAATTTCTGAGCCCTTGAGCTTCCCGATGATT 1789

DB 358 AGCTCACACACATTTTCCACAGAGATGCAATTTCTGAGCCCTTGAGCTTCCCGATGATT 299

QY 1790 CAAGAGAGAGAGTGGCAACAGGTTGGAGCTTATGAGCTTGACATCAGACTAT 1849

DB 298 CAGAGAGAGAGTGGCAACAGGTTGGAGCTTATGAGCTTGACATCAGACTAT 239

QY 1850 CGGTATCAGAAATAT 1865

DB 238 CGGTATCAGAAATAT 223

RESULT 12
LOCUS

BF197521/c 525 bp mRNA linear EST 03-NOV-2000
7084408.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3642903 3'

DEFINITION
6-O-SULFOTRANSFERASE. ; mRNA sequence.

BF197521
EST.

ACCESSION
VERSION
KEYWORDS

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
infoimage.llnl.gov
High quality sequence stop: 451.

FEATURES

SOURCE

Location/Qualifiers

1..525

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3642903"

/clone_lib="NCI-CGAP_Kid11"

/lab_host="DH10B"

/note="Organ: kidney; Vector: pRT3D-Pac (Pharmacia) with
a modified polylinker; Site: Not I; Site2: EcoRI;
Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonoids 1322376-1323911, 1456007-1456775, and
1500352-1502833). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT

ORIGIN

90 a 163 c 187 g 85 t

Query Match

2.9%; Score 59; DB 12; Length 525;

Best Local Similarity 100.0%; Pred. No. 9.1e-09;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGCTCTACAGCCAGCGTGGCTTCAAGAGAGTGGCTTTCACCTGACG 736

DB 260 AGGCGTGGCGCTCTACAGCCAGTGTGCTCAGAGAGTGGCTTTCACCTGACG 202

RESULT 13

LOCUS

AI824100/c 620 bp mRNA linear EST 21-DEC-1999
wJ46C01.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2405856 3'

DEFINITION

repetitive element ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert length: 917 Std Error: 0.00
Seq primer: -40bp from Gibco
High quality sequence stop: 490.

FEATURES

SOURCE

Location/Qualifiers

1..620

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="2405856"

/clone_lib="NCI-CGAP_Lu19"

/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"

/dev_stage="adult"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pRT3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pRT3D vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT

ORIGIN

106 a 195 c 227 g 89 t 3 others

Query Match

2.9%; Score 59; DB 9; Length 620;
Best Local Similarity 100.0%; Pred. No. 8e-09; Mismatches 0; Indels 0; Gaps 0;

Matches 59; Conservative 0;

QY 678 AGGCGTGGCGCTCTACAGCCAGCGTGGCTTCAAGAGAGTGGCTTTCACCTGACG 736

DB 314 AGGCGTGGCGCTCTACAGCCAGTGTGCTCAGAGAGTGGCTTTCACCTGACG 256

RESULT 14

BE857538/c

LOCUS	BE857538	695 bp	mRNA	linear	EST 29-SEP-2000
DEFINITION	7901a08.x1 NCI CGAP Brn23 Homo sapiens CDNA clone				IMAGE:3305174 3'
ACCESSION	BE857538				
VERSION	BE857538.1	GI:10371664			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 695)				
TITLE	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BRGP), Tumor Gene Index				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfield M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/TLNL, send email to: info@image.lnl.gov Seq primer: -40up from Gbpco High quality sequence stop: 431. Location/Qualifiers 1. 695 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3305174" /clone_lid="NCI CGAP_Brn23" /tissue_type="glioblastoma (poopled)" /lab_host="DH10B" /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st TGTACCATCTGTGAGTGGAGGGCGGCAATCTTTTTTTTTTTTTTTTTTTT T 3']: double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."				
FEATURES					
SOURCE					
BASE COUNT	127 a	210 c	247 g	111 t	
ORIGIN					
Query Match	2.98;	Score 59;	DB 12;	Length 695;	
Best Local Similarity	100.0%;	Pred. No. 7.4e-09;			
Matches 59;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 678	AGGCGCTGCGGCTCTTACAGCCAGCTGGTGGCTCAAGGAGGTGGCTTCTCAACCTGAG	736			
Db 272	AGGCGCTGCGGCTCTTACAGCCAGCTGGTGGCTCAAGGAGGTGGCTTCTCAACCTGAG	214			
LOCUS	BE858652	735 bp	mRNA	linear	EST 29-SEP-2000
DEFINITION	7501a09.x1 NCI CGAP Brn23 Homo sapiens CDNA clone				IMAGE:3305176 3'
ACCESSION	BE858652				
VERSION	BE858652.1	GI:10373890			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 735)				
TITLE	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BRGP), Tumor Gene Index				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfield M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/TLNL, send email to: info@image.lnl.gov Seq primer: -40up from Gbpco High quality sequence stop: 431. Location/Qualifiers 1. 735 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3305174" /clone_lid="NCI CGAP_Brn23" /tissue_type="glioblastoma (poopled)" /lab_host="DH10B" /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st TGTACCATCTGTGAGTGGAGGGCGGCAATCTTTTTTTTTTTTTTTTTTTT T 3']: double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."				

AUTHORS
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGA), Tumor Gene Index
JOURNAL COMMENT
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-rt@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LMN, send email to: info@lmn.lnl.gov
Seq primer: -400bp from Gibco
High quality sequence stop: 425.

FEATURES
source
Location/Qualifiers
1..735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3305176"
/clone_lib="NCI-CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I ; Site.2: Eco RI; 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCACTGTAGTGAGGGAGGCCGCATATCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
135 a 221 c 253 g 120 t 6 others

ORIGIN

Query Match 2.9%; Score 59; DB 12; Length 735;
Best Local Similarity 100.0%; Pred. No. 7.1e+09;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGCCGCTCCACAGCCAGCAGTCGTGATCAAGAGAGTGGCTTTCACACTGCAG 736
|||||
DB 328 AGGCTGCGCCGCTCCACAGCCAGCAGTCGTGATCAAGAGAGTGGCTTTCACACTGCAG 270

RESULT 16
AW081348/c
LOCUS
DEFINITION
Accession
 accession
KEYWORDS
SOURCE
ORGANISM
human.
 Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 436)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-rt@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/dbp/image.html

Seq primer: -400P from gldco
High quality sequence stop: 431.

FEATURES

source

1. 436

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1ib="NCI_CGAP_Co20"

/issue_type="moderately differentiated adenocarcinoma"

/lab_host="DH10B"

/note="Organ: colon; Vector: PCMV-SPORT6; Site: 1; Salt: Site: 2; Not: Cloned unidirectionally. Primer: Oligo dt. Normalized to Cot 500. Average insert size 1.1kb. Normalized version of NCI_CGAP_Co18. Library constructed by Life Technologies."

BASE COUNT

64 a 136 c 169 g 67 t

ORIGIN

Query Match 2.3%; Score 47; DB 10; Length 436;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 678 AGGCGTCCGCTCTACAGCCAGTGTGCTCAAGAGTGCCTTC 724
|||||
DB 309 AGGCGTCCGCTCTACAGCCAGTGTGCTCAAGAGTGCCTTC 263

RESULT 17
BM246681
LOCUS
DEFINITION
K0741E04-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
CDNA Library (Long) Mus musculus CDNA clone K0741E04 3', mRNA
sequence.
ACCESSION
BM246681
VERSION
BM246681.1
KEYWORDS
GI:17881951
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 497)

REFERENCE
AUTHORS
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Alba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)
Unpublished (2001)
CONTACT: Dawood B. Dudekula
LABORATORY OF GENETICS
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0741 row: E column: 04
Seq primer: -21M13 forward
High quality sequence stop: 497
POLY-A-yes.

FEATURES

source

Location/Qualifiers

1. 497

/organism="Mus musculus"

/strain="C57BL/6NCR"

/db_xref="nlaEST:K0741E04-3"

/db_xref="taxon:10090"

/clone_1ib="K0741E04"

/clone_1ib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)"

/tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-
)"

/dev_stage="Age approx. 10 weeks old"

/lab_host="DH10B"

/note="Vector: pSPORT1 (Invitrogen); Site: 1; Salt: Site: 2;
Not: Mouse CDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research

BASE COUNT

87 a 137 c 134 g 139 t

ORIGIN

Query Match 1.3%; Score 27; DB 13; Length 497;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 494 GACATGAGCGCTCTTGATGCTACATG 520
|||||
DB 441 GACATGAGCGCTCTTGATGCTACATG 467

RESULT 18
BM245312
LOCUS
DEFINITION
BM245312
K0722H07-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
CDNA Library (Long) Mus musculus CDNA clone K0722H07 3', mRNA
sequence.
ACCESSION
BM245312
VERSION
BM245312.1
KEYWORDS
GI:17880582
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 536)

REFERENCE
AUTHORS
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Alba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)
Unpublished (2001)
CONTACT: Dawood B. Dudekula
LABORATORY OF GENETICS
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0722 row: H column: 07
Seq primer: -21M13 forward
High quality sequence stop: 536
POLY-A-yes.

FEATURES

source

Location/Qualifiers

1. 536

/organism="Mus musculus"

/strain="C57BL/6NCR"

/db_xref="nlaEST:K0722H07-3"

/db_xref="taxon:10090"

/clone_1ib="K0722H07"

/clone_1ib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)"

/tissue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-
)"

/dev_stage="Age approx. 10 weeks old"

```

/lab_host="DH10B"
/Note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://19sun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCCTAGATCGGAGCGGCCCTTTTCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker IL-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Tag polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."

BASE COUNT      93 a      150 c      147 g      146 t
ORIGIN

Query Match      1.3%; Score 27; DB 13; Length 536;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCTACATG 520
|||||
Db 441 GACATGAGCGCTTTGATGCTACATG 467

RESULT 19
BG9664671 695 bp mRNA linear EST 12-JUN-2001
LOCUS 602831875F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4986258 5',
DEFINITION mRNA sequence.
ACCESSION BG9664671
VERSION BG9664671.1 GI:14352308
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 695)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM10995 row: f column: 19
High quality sequence stop: 684.
Location/Qualifiers
1..695
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4986258"
/clone_1id="NCI_CGAP_Co24"
/lab_host="DH10B (TI phage-resistant)"
/Note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;

/lab_host="DH10B"
/Note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://19sun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCCTAGATCGGAGCGGCCCTTTTCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker IL-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Tag polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."

BASE COUNT      93 a      150 c      147 g      146 t
ORIGIN

Query Match      1.3%; Score 27; DB 13; Length 536;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCTACATG 520
|||||
Db 441 GACATGAGCGCTTTGATGCTACATG 467

RESULT 19
BG9664671 695 bp mRNA linear EST 12-JUN-2001
LOCUS 602831875F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4986258 5',
DEFINITION mRNA sequence.
ACCESSION BG9664671
VERSION BG9664671.1 GI:14352308
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 695)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM10995 row: f column: 04
High quality sequence stop: 741.
Location/Qualifiers
1..783
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4982643"
/clone_1id="NCI_CGAP_Co24"
/lab_host="DH10B (TI phage-resistant)"
/Note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT      141 a      236 c      226 g      180 t
ORIGIN

Query Match      1.3%; Score 27; DB 13; Length 783;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCTACATG 520
|||||
Db 376 GACATGAGCGCTTTGATGCTACATG 402

RESULT 21
BG966340 852 bp mRNA linear EST 12-JUN-2001
LOCUS 60283286F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4987327 5',
DEFINITION mRNA sequence.
ACCESSION BG966340
VERSION BG966340.1 GI:14353977

```


CAGAGAGATTCGAGCTAATTAATCCCGCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

Location/Qualifiers
1..1923

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2310003618"
/db_xref="MGI:1904558"
/db_xref="taxon:10090"
/clone="2310003618"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1923
/gene="Chst4"
96..1262
/gene="Chst4"
/note="carbohydrate (chondroitin 6/keratan)
sulfotransferase 4
data source:MGI, source key:MGI:1349479, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB26078.1"
/db_xref="GI:12843702"
/db_xref="MGI:1349479"
/translation="MMLIKKGRLLMFLGQVYVVALTHNSVHNLQRESRRPVY
LVLSMRSGSSVFQQLFQHPDVYLYEPANHWMTTSTAMKLNHNAVRLNSVL
CDMSVFAYNMPGRKOSLFEQMSALCAPCDEFPAREHISPPHCKILCCQEP
DMVKACSHGFVYLVKEVRLPLQALYPLDPLNLHVVLVADPRAVFRSEHTT
ELMVDSHVILQDLLETKIEEDQPYAMKTIICKSOVDYKAIQTLPEALQCYLRE
DLVAPLAQOTRLKFGVLDPLPLHLOTVMYVTEGKGGMGHAFHFNRLNLYQAR
MSLVEKYSQIQDQCGEMDLGLGYRSGQEGCNLSIDLSSHILGVFRFG"

gene
CDS

BASE COUNT 445 a 533 c 492 g 453 t
ORIGIN

Query Match

Best Local Similarity 1.3%; Score 27; DB 11; Length 1923;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTGTGATGCTCATG 520
|||||
Db 405 GACATGAGCGCTTGTGATGCTCATG 431

RESULT 23
BM036473/c 501 bp mRNA linear EST 05-NOV-2001
LOCUS fub1f07.x1 Gong zebrafish testis Danio rerio cDNA clone 5377309 3',
DEFINITION mRNA sequence.
ACCESSION BM036473
VERSION BM036473.1 GI:16750044
KEYWORDS EST.
SOURCE zebrafish
ORGANISM Danio rerio

REFERENCE

1 (bases 1 to 501)
Clark,M., Johnson,S.L., Lebrach,H., Lee,R., Li,F., Marra,M., Eddy,
'S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
'K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritzer,E.,
Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.

TITLE

Washu zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES

Location/Qualifiers
1..501

/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5377309"
/clone_lib="Gong zebrafish testis"
/sex="male"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis (pooled); Vector: pBluescript SK-;
Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
the testes of 31 male adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excision to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."

BASE COUNT 171 a 94 c 90 g 146 t
ORIGIN

Query Match

Best Local Similarity 1.2%; Score 25; DB 13; Length 501;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 CCCAATATAGAAAAA 2032
|||||
Db 29 CCCAATATAGAAAAA 5

RESULT 24
CNS02BKQ/c 682 bp DNA linear GSS 12-MAY-2000
LOCUS Tetradon nigroviridis genome survey sequence pUC-ori end of clone
DEFINITION 253N23 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL189971
VERSION AL189971.1 GI:7828075
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis

REFERENCE

1 (bases 1 to 682)
Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Barnot,A., Fitzmes,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.

TITLE

Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished

REFERENCE

2 (bases 1 to 682)
Roest-Crolius,H., Jallion,O., Dasilva,C., Fitzmes,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.

TITLE

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis

REFERENCE

3 (bases 1 to 682)
Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (12-APR-2000)

COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

Location/Qualifiers

1..682

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="253N23"

/clone.lib="g"

/note="Tetraodon sequence ID : C0AG253CG12SP1-end : PUC-ori"

BASE COUNT 202 a 122 c 113 g 224 t 21 others

ORIGIN

Query Match 1.2% ; Score 25; DB 17; Length 682;

Best Local Similarity 100.0% ; Pred. No. 2.9e+02;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 CCCCAATAGAGAAAAAAAAAAAAA 2032

Db 241 CCCCAATAGAGAAAAAAAAAAAAA 217

Search completed: February 22, 2003, 22:52:05
Job time : 2651 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 19:04:25 ; Search time 5242 Seconds

(without alignments)
11281.362 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032

Sequence: 1 ggcctcgagccagagatgct.....ataagaaaaaaaaaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 25

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl:*

- 1: gb_da:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_pu:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_hlg_hum:*
- 31: em_hlg_inv:*
- 32: em_hlg_other:*
- 33: em_hlg_mus:*
- 34: em_hlg_pln:*
- 35: em_hlg_rtd:*
- 36: em_hlg_mam:*
- 37: em_hlg_vrl:*
- 38: em_sy:*
- 39: em_hlg_hum:*
- 40: em_hlg_mus:*
- 41: em_hlg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	6 AR203335	AR203335 Sequence
2	2032	100.0	2032	9 AF131235	AF131235 Homo sapi
3	1856	91.3	2011	9 AK026635	AK026635 Homo sapi
4	1806	88.9	183228	9 AC010547	AC010547 Homo sapi
5	1756	86.4	1992	9 AF280088	AF280088 Homo sapi
6	1211	59.6	1333	9 AF149783	AF149783 Homo sapi
7	505	24.9	517	6 AX381256	AX381256 Sequence
8	286	14.1	389	6 AX381326	AX381326 Sequence
9	59	2.9	1462	6 AF176839	AF176839 Homo sapi
10	59	2.9	1647	6 AX327330	AX327330 Sequence
11	59	2.9	2170	9 AF176838	AF176838 Homo sapi
12	59	2.9	2544	9 AF219990	AF219990 Homo sapi
13	59	2.9	3278	9 AF246718	AF246718 Homo sapi
14	59	2.9	3786	9 AF280086	AF280086 Homo sapi
15	59	2.9	71503	9 AF219991	AF219991 Homo sapi
16	59	2.9	157337	9 AC009163	AC009163 Homo sapi
17	59	2.9	194832	2 AC025287	AC025287 Homo sapi
18	59	2.9	208185	2 AC009105	AC009105 Homo sapi
19	37	1.8	37	6 AR203340	AR203340 Sequence
20	27	1.3	1926	10 AF131236	AF131236 Mus muscu
21	27	1.3	2201	10 AF109155	AF109155 Mus muscu

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	SEQUENCE 2	FROM PATENT	US 6365365
AR203335	AR203335	Sequence 2 from patent US 6365365.			
AR203335	AR203335	Sequence 2 from patent US 6365365.			
AR203335.1	GI:21499698				
KEYWORDS	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 2032)				
AUTHORS	Bistrup A., Rosen, S.D., Tangemann, K. and Hemmerich, S.				
TITLE	Method of determining whether an agent modulates glycosyl				
JOURNAL	sulfotransferase-3				
FEATURES	Patent: US 6365365-A 2 02-APR-2002;				
SOURCE	Location/Qualifiers				
1. 2032					
BASE COUNT	468 a 569 c 490 g 505 t				
ORIGIN	/organism="unknown"				
Query Match	100.0%; Score 2032; DB 6; Length 2032;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 2032; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 GGCCTCGAGCCAGAGATCCCTCCAGTCTGGGGGAAAAATGCTTCTCTCAATTCCTCTCCAG 60				
DB	1 GGCCTCGAGCCAGAGATCCCTCCAGTCTGGGGGAAAAATGCTTCTCTCAATTCCTCTCCAG 60				
QY	61 CCCACCTCAAGCAGTCTCCACCACCCCTTGAGTCTCAAGCAGTGTAAAGCGTACTTCA 120				
DB	61 CCCACCTCAAGCAGTCTCCACCACCCCTTGAGTCTCAAGCAGTGTAAAGCGTACTTCA 120				
QY	121 CAGCTTCTCGGAGCAGAGTCTTCTCAAGCCGCTCTTCAAGGCTTCCACTTCAGCAG 180				
DB	121 CAGCTTCTCGGAGCAGAGTCTTCTCAAGCCGCTCTTCAAGGCTTCCACTTCAGCAG 180				
QY	181 AATGCTACTGCTTAAAAAATGAAGCTCCGCTGTTCTGCTTCCAGATGCCATCTT 240				

[illegible]

TITLE		Hemmerich, S.	
JOURNAL		Direct Submission	
FEATURES		Submitted (25-FEB-1999) Respiratory Diseases, Roche Bioscience, 3401 Hillview Avenue, Palo Alto, CA 94304, USA	
source		Location/Qualifiers	
		1..2032	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/cell_type="high endothelial"	
		/tissue_type="tonsil"	
		182..1342	
CDS		/function="biosynthesis of endothelial ligands for L-selection"	
		/note="HEC-G1CMAC-6-ST"	
		/codon_start=1	
		/product="N-acetylglucosamine 6-O-sulfotransferase"	
		/protein_id="AAD3015.1"	
		/db_xref="GI:4927114"	
		/translation="MLPKKKKLLFLVSQMAILLFFHMYSHNISLSMKAOPERMH VLVLSMRSGSEVGOLEFGHPDVEYLMPEAMVHMTFQKOSTAMLMHAYRDLIRAY LCDSVFAYMEPGPBRSSSLFQWENSALCASAPACILIPODELIPRAHRLISQOP FEVEKARSRYSHVILKEVREPNQSLXPILKPSLNLHLYLDRPRAVRSERK GDLMIDRIYVGHQKIKREDOPYVMQVTCOSOLETKITQSLPKALDERLYLVAY EDLRAFAVQTSRMTEFVGLFELPHLQTVHNTIRGMDHAFETNARALANSQAM RWSLPYEVKSLQKACGDANMLLGRVHRSEQRNLDLSTWIVPEQIH"	
BASE COUNT		468 a 569 c 490 g 505 t	
ORIGIN			
Query Match		100.0%; Score 2032; DB 9; Length 2032;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 2032; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GGCTTGAGCGCAGATGCTCCAGTCTGTGGGGAAAATGCTTCCTCATTTGCTTCCAG	60
DB	1	GGCTGAGCGCAGATGCTCCAGTCTGTGGGGAAAATGCTTCCTCATTTGCTTCCAG	60
QY	61	CCACCTCAAGCAGTCTCCACCCCTTGAGTCTCAGCAGTGAAGCTGTACTTCA	120
DB	61	CCACCTCAAGCAGTCTCCACCCCTTGAGTCTCAGCAGTGAAGCTGTACTTCA	120
QY	121	CAGCTTCTGGGAGGAGTGTCTTCTCAAGCCGCTGTCAAGGCTCTCCACTCAGAC	180
DB	121	CAGCTTCTGGGAGGAGTGTCTTCTCAAGCCGCTGTCAAGGCTCTCCACTCAGAC	180
QY	181	AATGCTACTGCTAAAAAATGAGCTCTGCTGTTTCCAGATGCGCATCTT	240
DB	181	AATGCTACTGCTAAAAAATGAGCTCTGCTGTTTCCAGATGCGCATCTT	240
QY	241	GGCTTATTTCTCCAGATGACAGCACAATCAGTCCCTGCTATGAGAGCAGACCC	300
DB	241	GGCTTATTTCTCCAGATGACAGCACAATCAGTCCCTGCTATGAGAGCAGACCC	300
QY	301	CGAGCGATGACGTGTGTTCTCTTCTCTGCGGCTCTGCTGCTTCTTTTGTGGGCA	360
DB	301	CGAGCGATGACGTGTGTTCTCTTCTCTGCGGCTCTGCTGCTTCTTTTGTGGGCA	360
QY	361	GCTTTTGGGAGACCCAGATGTTTCTTACTGATGAGACCCGCTGGACGCTGTGAT	420
DB	361	GCTTTTGGGAGACCCAGATGTTTCTTACTGATGAGACCCGCTGGACGCTGTGAT	420
QY	421	GACCTTCAGAGACACCGCTGATGCTCAGATGCTGTGGGGATCTGATACGGGC	480
DB	421	GACCTTCAGAGACACCGCTGATGCTCAGATGCTGTGGGGATCTGATACGGGC	480
QY	481	CGTCTTCTGTGCGACATGAGCTTTGATGCTACATGGAACCTGTCTCCCGGAGCA	540
DB	481	CGTCTTCTGTGCGACATGAGCTTTGATGCTACATGGAACCTGTCTCCCGGAGCA	540
QY	541	GTCACGCTCTTTAGTGGAGAACAGCGGCTGTGTCTGACCTGCTGTGACAT	600
DB	541	GTCACGCTCTTTAGTGGAGAACAGCGGCTGTGTCTGACCTGCTGTGACAT	600
QY	601	CATCCCAAGATGAATCATCCCCCGGCTCACTGACAGGCTCCTGTGACATCAAGCC	660
DB	601	CATCCCAAGATGAATCATCCCCCGGCTCACTGACAGGCTCCTGTGACATCAAGCC	660
QY	661	CTTGAGCTGTGAGAGGCTGCGGCTCTTACAGCCACGTGTGCTCAAGAGGTGCG	720
DB	661	CTTGAGCTGTGAGAGGCTGCGGCTCTTACAGCCACGTGTGCTCAAGAGGTGCG	720
QY	721	CTTCTCACTGACAGTCCCTCTACCCGCTGCTGAAGACCCCTCCCTCAACTGATAT	780
DB	721	CTTCTCACTGACAGTCCCTCTACCCGCTGCTGAAGACCCCTCCCTCAACTGATAT	780
QY	781	CGTGACCTGTGCGGGAGACCCCGGCGCGTCTCCGTTCCGSAAGACCAAGGAGA	840
DB	781	CGTGACCTGTGCGGGAGACCCCGGCGCGTCTCCGTTCCGTTCCGSAAGACCAAGGAGA	840
QY	841	TCTCATGATTGACAGTGCATTGTGATGGGACAGCATGAGCAGAACTCAAGAGAGA	900
DB	841	TCTCATGATTGACAGTGCATTGTGATGGGACAGCATGAGCAGAACTCAAGAGAGA	900
QY	901	CCAACCCCTACTATGTGATGAGAGGTGATCTGCAAGCCAGCTGAGATCTTCAAGACAT	960
DB	901	CCAACCCCTACTATGTGATGAGAGGTGATCTGCAAGCCAGCTGAGATCTTCAAGACAT	960
QY	961	CCAGTCTGTGCCAAGGCGCTGACAGAAAGCTGCTGCTGTGCGCTATGAGACCTGCG	1020
DB	961	CCAGTCTGTGCCAAGGCGCTGACAGAAAGCTGCTGCTGTGCGCTATGAGACCTGCG	1020
QY	1021	TGAGACCCCTGTGCCCCAGACTCCCGAATGATGAATTGCTGGATGGAATTTTGGC	1080
DB	1021	TGAGACCCCTGTGCCCCAGACTCCCGAATGATGAATTGCTGGATGGAATTTTGGC	1080
QY	1081	CCATCTTCAGACCTGGGGGCAATACATCACCCGAGGCAATGGGTACACCGCTTT	1140
DB	1081	CCATCTTCAGACCTGGGGGCAATACATCACCCGAGGCAATGGGTACACCGCTTT	1140
QY	1141	CCACCAAAATGCCAGGAGTGCCTTAATGTCTCCAGAGCTGTGGCGTGTGGCCCTA	1200
DB	1141	CCACCAAAATGCCAGGAGTGCCTTAATGTCTCCAGAGCTGTGGCGTGTGGCCCTA	1200
QY	1201	TGAAAGGTTTCTGCACTTCAAGAAAGCTGTGGGATGCCATGATTTGCTGGGCTACG	1260
DB	1201	TGAAAGGTTTCTGCACTTCAAGAAAGCTGTGGGATGCCATGATTTGCTGGGCTACG	1260
QY	1261	CCAGTCAAGATCTGAACAGAGAAAGAACTGCTGATCTTCTGCTACCTGAC	1320
DB	1261	CCAGTCAAGATCTGAACAGAGAAAGAACTGCTGATCTTCTGCTACCTGAC	1320
QY	1321	TGTCCCTGAGCAATTCACCTAAGAGGTTGAGAGGCTTTGCTGCCACCTGTGTACGCC	1380
DB	1321	TGTCCCTGAGCAATTCACCTAAGAGGTTGAGAGGCTTTGCTGCCACCTGTGTACGCC	1380
QY	1381	TCAGTCACTTCTCTGATGCTTCTGAGCCTTGCTCTCATCTGAGCCTTAATCATATG	1440
DB	1381	TCAGTCACTTCTCTGATGCTTCTGAGCCTTGCTCTCATCTGAGCCTTAATCATATG	1440
QY	1441	TCGTGGGTATCACACTGAGGTGAGTTGTGTCCACAGCTGTCAACAGAGACTTTT	1500
DB	1441	TCGTGGGTATCACACTGAGGTGAGTTGTGTCCACAGCTGTCAACAGAGACTTTT	1500
QY	1501	GTCGCAATGCTGTGTAGAGAAACAGAGGGAACCTATATGAGCAGACATCCAC	1560
DB	1501	GTCGCAATGCTGTGTAGAGAAACAGAGGGAACCTATATGAGCAGACATCCAC	1560
QY	1561	CAGTGAAGAGGGATGCTCTTCTTCTTCTGATCTTCTGCTGTGGCAGACTTCA	1620
DB	1561	CAGTGAAGAGGGATGCTCTTCTTCTTCTGATCTTCTGCTGTGGCAGACTTCA	1620
QY	1621	AGACTTTGTGGCGTGGAGGCTATTAAGCAGACAGAGTATGAGTGAATGATCATAA	1680
DB	1621	AGACTTTGTGGCGTGGAGGCTATTAAGCAGACAGAGTATGAGTGAATGATCATAA	1680
QY	1681	ACCTCCCTGTCAATCTTCCCAATGGGAGATGATCTTCCAAAGAGCTCACACAG	1740
DB	1681	ACCTCCCTGTCAATCTTCCCAATGGGAGATGATCTTCCAAAGAGCTCACACAG	1740

Db 1681 ACCTCCCTGTCACATCTTGGCCCAATGGGGAATGATCTTTTCACCAAGAGCTCCACACG 1740

QY 1741 ATTTTCCACAGAGATCAATTTCTGAGCCCTGGAGTCCAGTGATTCAGGAAGGA 1800

Db 1741 ATTTTCCACAGAGATCAATTTCTGAGCCCTGGAGTCCAGTGATTCAGGAAGGA 1800

QY 1801 GTGGGAACAAGTTGGATGCTTATGAGCTTGACCAATCAGGATTCGGTAATCAGA 1860

Db 1801 GTGGGAACAAGTTGGATGCTTATGAGCTTGACCAATCAGGATTCGGTAATCAGA 1860

QY 1861 AATATGAACAAATCTCTGACACAAAGCAAGCTCTTAAGTTCAAGGGTGGCTGGGC 1920

Db 1861 AATATGAACAAATCTCTGACACAAAGCAAGCTCTTAAGTTCAAGGGTGGCTGGGC 1920

QY 1921 TGCATTGATATCACTTCCCTCTGATTTTCCCATCATGAGAGAGCTTGGAGTGG 1980

Db 1921 TGCATTGATATCACTTCCCTCTGATTTTCCCATCATGAGAGAGCTTGGAGTGG 1980

QY 1981 AAGTGGCATCTGTATATCTAAATTTCCCAATTAAGAAAAA 2032

Db 1981 AAGTGGCATCTGTATATCTAAATTTCCCAATTAAGAAAAA 2032

RESULT 3

AK026635 2011 bp mRNA linear PRI 29-SEP-2000

LOCUS AK026635

DEFINITION Homo sapiens cDNA: FLJ22982 fls, clone KAT11454, highly similar to AF131235 Homo sapiens N-acetylglyucosamine 6-O-sulfotransferase

MRNA

ACCESSION AK026635 GI:10439531

VERSION AK026635.1

KEYWORDS oligo capping; fls (full insert sequence).

SOURCE Homo sapiens signal-ring cell carcinoma cell_line:KATO III cDNA to MRNA, clone 11b:KAT clone:KAT11454.

ORGANISM Homo sapiens

REFERENCE Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Ohashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

TITLE NEDO human cDNA sequencing project

JOURNAL unpublished

REFERENCE 2 (bases 1 to 2011)

AUTHORS Sugano,S., Suzuki,Y., Ota,T., Ohashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - 3' -end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. .2011

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="KAT11454"

/cell_line="KATO III"

/cell_type="signal-ring cell carcinoma"

/clone_lib="KAT"

/note="cloning vector pME18SPL3"

misc_feature

1. .2011

/note="highly similar to AF131235 Homo sapiens N-acetylglyucosamine 6-O-sulfotransferase mRNA"

BASE COUNT 482 a 541 c 475 g 513 t

Query Match 91.3%; Score 1856; DB 9; Length 2011;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 111 GTTACTTTACAGCTTCTCTGGAGGAGAGTGTCTTCTCAAGCCCGCTTGAAGGTCTTC 170

Db 1 GTTACTTTACAGCTTCTCTGGAGGAGAGTGTCTTCTCAAGCCCGCTTGAAGGTCTTC 60

QY 171 ACTTAGACATGCTACTGGCTTAAAAATGAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 230

Db 61 ACTTAGACATGCTACTGGCTTAAAAATGAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 120

QY 231 TGGCCATCTTGGCTCTATCTTCTCAAGATGATACAGCCACACATCAGCTCCCTCTATGA 290

Db 121 TGGCCATCTTGGCTCTATCTTCTCAAGATGATACAGCCACACATCAGCTCCCTCTATGA 180

QY 291 AGGCACAGCCCGAGCGCATGACAGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 350

Db 181 AGGCACAGCCCGAGCGCATGACAGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 240

QY 351 TTGTGGGAGAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGAGCCCGCTGGC 410

Db 241 TTGTGGGAGAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGAGCCCGCTGGC 300

QY 411 ACGTGTGATGACCTTCAAGCAGACACCGCTGATGCTGACATGCTGTGGGATC 470

Db 301 ACGTGTGATGACCTTCAAGCAGACACCGCTGATGCTGACATGCTGTGGGATC 360

QY 471 TGATACGGGGCGCTTCTTGTGCGACATGAGAGCTTTTGATGCTTACATGAGACCTGTG 530

Db 361 TGATACGGGGCGCTTCTTGTGCGACATGAGAGCTTTTGATGCTTACATGAGACCTGTG 420

QY 531 CCCGAGACAGTCCAGCTCTTCAAGTGGAGAACAGCCGGGCGCTGTGTCTGACCTG 590

Db 421 CCCGAGACAGTCCAGCTCTTCAAGTGGAGAACAGCCGGGCGCTGTGTCTGACCTG 480

QY 591 CCTGTGACATCATCCCAAGATGAAATCATCCCGGGCGCTACAGCAGGCTCGTGCA 650

Db 481 CCTGTGACATCATCCCAAGATGAAATCATCCCGGGCGCTACAGCAGGCTCGTGCA 540

QY 651 GTCAACAGCCCTTATGAGTGTGGAGAGAGCCCTGCGCTCTACAGCCAGCTGTGCTCA 710

Db 541 GTCAACAGCCCTTATGAGTGTGGAGAGAGCCCTGCGCTCTACAGCCAGCTGTGCTCA 600

QY 711 AGGAGTGGCTTCTTCAAGCTGCAAGCTCTTACCCTGCTGTAAGAACCCCTCCCTCA 770

Db 601 AGGAGTGGCTTCTTCAAGCTGCAAGCTCTTACCCTGCTGTAAGAACCCCTCCCTCA 660

QY 771 ACGTGCATCTGTGACCTGCTCCGGAGCCCGGGCGCTGTCCGTTCCGAGAACGCA 830

Db 661 ACGTGCATCTGTGACCTGCTCCGGAGCCCGGGCGCTGTCCGTTCCGAGAACGCA 720

QY 831 CAAGAGAGATCTCATGATTTGACAGTCCGATTTGATGGGAGCATGAGCAAGAACTCA 890

Db 721 CAAGAGAGATCTCATGATTTGACAGTCCGATTTGATGGGAGCATGAGCAAGAACTCA 780

QY 891 AGAAGGAGACCAACCCCTCTATGTGATGATGATGATGATGATGATGATGATGATGAT 950

Db 781 AGAAGGAGACCAACCCCTCTATGTGATGATGATGATGATGATGATGATGATGATGAT 840

QY 951 ACAAGACATCTCAAGCTTGGCCAAAGGCCCTCAGAGAACGATCTGTGGCTATG 1010

Db 841 ACAAGACATCTCAAGCTTGGCCAAAGGCCCTCAGAGAACGATCTGTGGCTATG 900

QY 1011 AGGACCTGGCTGAGCCCTGTGGCCCAAGCTTCCCGAATGATATGATGATGATGATG 1070

Db 901 AGGACCTGGCTGAGCCCTGTGGCCCAAGCTTCCCGAATGATATGATGATGATGATG 960

QY 1071 AATCTTGGCCATCTTCAAGCTGGTGGATGATGATGATGATGATGATGATGATGATGAT 1130

Db 961 AATCTTGGCCATCTTCAAGCTGGTGGATGATGATGATGATGATGATGATGATGATGAT 1020

QY 1131 ACCAGCCTTCCACACAAATGCCAGGATGCCCTAATGCTCCAGGCTGGCGTGT 1190
 Db 1021 ACCAGCCTTCCACACAAATGCCAGGATGCCCTAATGCTCCAGGCTGGCGTGT 1080
 QY 1191 CTTTGCCCATGAAAAGTTCTGCATCTCAAAAGCCTGGGCAATGCCAATTTGC 1250
 Db 1081 CTTTGCCCATGAAAAGTTCTGCATCTCAAAAGCCTGGGCAATGCCAATTTGC 1140
 QY 1251 TGGGCTACCGCCACGCTGAGATCTGAACAAGACAGAAACCTGTGCTGATCTTGT 1310
 Db 1141 TGGGCTACCGCCACGCTGAGATCTGAACAAGACAGAAACCTGTGCTGATCTTGT 1200
 QY 1311 CTAAGTGAAGCTGCTGAGAGCAAACTCACTAAGAGGTTGAGAGGCTTTGCTGCACCT 1370
 Db 1201 CTAAGTGAAGCTGCTGAGAGCAAACTCACTAAGAGGTTGAGAGGCTTTGCTGCACCT 1260
 QY 1371 GGTGTCAGCCTCAGTCACTCTTCTGTGAATGCTTGTGAGCCTTGTGCTGATCTGAGCCT 1430
 Db 1261 GGTGTCAGCCTCAGTCACTCTTCTGTGAATGCTTGTGAGCCTTGTGCTGATCTGAGCCT 1320
 QY 1431 TAATTAATGCTGTGCTGATCACTGAGTGTGATGCTGTGCAACAGCTGCTCAAGCAG 1490
 Db 1321 TAATTAATGCTGTGCTGATCACTGAGTGTGATGCTGTGCAACAGCTGCTCAAGCAG 1380
 QY 1491 AAGGACTTTGTGCTGATGCTGCTGTGCTAAGAAACAGCTGGGAACTTATGTGAGCAG 1550
 Db 1381 AAGGACTTTGTGCTGATGCTGCTGTGCTAAGAAACAGCTGGGAACTTATGTGAGCAG 1440
 QY 1551 CACATCCACAGTGAAGAAACAGGATGCTGCTGCTTCTTCTGATCTTCTGCTGCTGCTGCTG 1610
 Db 1441 CACATCCACAGTGAAGAAACAGGATGCTGCTGCTTCTTCTGATCTTCTGCTGCTGCTGCTG 1500
 QY 1611 CAGACTTCAGAGACTTTGTGCTGCTGAGAGCCTATTAAGCAGACAGATGAGTGAAT 1670
 Db 1501 CAGACTTCAGAGACTTTGTGCTGCTGAGAGCCTATTAAGCAGACAGATGAGTGAAT 1560
 QY 1671 TGATCCATTAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1730
 Db 1561 TGATCCATTAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 QY 1731 GCTACCCAGCATTTTCCACAGATGCAAAATTCGAGGCCCTGAGAGTCCAGTGAATTC 1790
 Db 1621 GCTACCCAGCATTTTCCACAGATGCAAAATTCGAGGCCCTGAGAGTCCAGTGAATTC 1680
 QY 1791 AAGGAGGAATGCGGAACAAAGTTGAGTGCCTACTTAATGAGCTGAGCATCAAGCTATC 1850
 Db 1681 AAGGAGGAATGCGGAACAAAGTTGAGTGCCTACTTAATGAGCTGAGCATCAAGCTATC 1740
 QY 1851 GGTATATCAAAATATGAAAACAAATCTCTGACAAAAGAGCAAGCTCTTAAGTTCACAGG 1910
 Db 1741 GGTATATCAAAATATGAAAACAAATCTCTGACAAAAGAGCAAGCTCTTAAGTTCACAGG 1800
 QY 1911 GTGCTGGGGCGCATTTGAATATGATCTCCCTGCTGATTTTCCATCAATAGAAACT 1970
 Db 1801 GTGCTGGGGCGCATTTGAATATGATCTCCCTGCTGATTTTCCATCAATAGAAACT 1860
 QY 1971 TTGACCTGTGAAGCTGCATCTGTTAATAGTAAATTCGCAATTAAG 2017
 Db 1861 TTGACCTGTGAAGCTGCATCTGTTAATAGTAAATTCGCAATTAAG 1907

RESULT 4
 AC010547 183228 bp DNA linear PRI 29-SEP-2001
 LOCUS Homo sapiens chromosome 16 clone RP11-510M2, complete sequence.
 DEFINITION AC010547
 ACCESSION AC010547
 VERSION AC010547.9 GI:15808510
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 183228)

AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 183228)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 183228)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (29-SEP-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT On Sep 29, 2001 this sequence version replaced gi:14589436.
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

FEATURES
 source 1.183228
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-510M2"
 BASE COUNT 51901 a 42098 c 40026 g 49203 t
 ORIGIN

Query Match 88.9%; Score 1806; DB 9; Length 183228;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1856; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 161 AAGTCTTCACCTTCAAGCAATGCTACTGCTAAAAATGAAAGTCTGCTGTTTCG 220
 Db 142584 AAGTCTTCACCTTCAAGCAATGCTACTGCTAAAAATGAAAGTCTGCTGTTTCG 142643
 QY 221 GTTTCCAGATGCGCATCTTGCTCTATTCTTCCATGATGACGCCACAATCACTCGC 280
 Db 142644 GTTCCAGATGCGCATCTTGCTCTATTCTTCCATGATGACGCCACAATCACTCGC 142703
 QY 281 CTGCTATGAAGGACAGACCCGCGGCGCATGACAGTGTGCTGCTTCTGCGCTCT 340
 Db 142704 CTGCTATGAAGGACAGACCCGCGGCGCATGACAGTGTGCTGCTTCTGCGCTCT 142763
 QY 341 GGCCTCTCTTTTGTGGGGAGCTTTTGGGACAGCACCAGATGTTTCTACCGATGAG 400
 Db 142764 GGCCTCTCTTTTGTGGGGAGCTTTTGGGACAGCACCAGATGTTTCTACCGATGAG 142823
 QY 401 CCGGCTGGAGCTGTGATGACCTTCAAGCAGACCGCGCTGATGCTGACATGCT 460
 Db 142824 CCGGCTGGAGCTGTGATGACCTTCAAGCAGACCGCGCTGATGCTGACATGCT 142883
 QY 461 GTGCGGATCTGATGAGGCGCGCTCTTGTGAGACATGAGGCTTTGATGCTATCATG 520
 Db 142884 GTGCGGATCTGATGAGGCGCGCTCTTGTGAGACATGAGGCTTTGATGCTATCATG 142943
 QY 521 GAACCTGTGCTCCCGGAGACATGCTTCAAGTCTTCAAGTCTGAGAGAACCGCGCTGTGT 580

OY	1661	TCAGGCAATTGATCCATAAACCCTGCCGTGCACATCTTTGGCCAAATGGGAATGATCTT	1720
Db	144084	TCAGTGGAAATTGATCCATAAACCCTGCCGTGCACATCTTTGGCCAAATGGGAATGATCTT	144143
OY	1721	TCACCAAGAGCTCACCGCATTTTCCACAGAGATGCAAATTTCTGAGCCCTTGAGGTCC	1780
Db	144144	TCACCAAGAGACTCACCAGCATTTTCCACAGAGATGCAAATTTCTGAGCCCTTGAGGTCC	144203
OY	1781	CAGTGCATTCAAGSAGAAGAGTGGGACAAGATTGATGAGCTACTTATGAGCTTGACCAT	1840
Db	144204	CAGTGCATTCAAGSAGAAGAGTGGGACAAGATTGATGAGCTACTTATGAGCTTGACCAT	144263
OY	1841	CACACCTATCGGTATCAATCAAAAATATGAAAAACAATAATCTCTGCACAAAAAAGCAAGTCTTA	1900
Db	144264	CACACCTATCGGTATCAAGAAATATGAAAAACAATAATCTCTGCACAAAAAAGCAAGTCTTA	144323
OY	1901	AGTTACAGAGGGGCGCTGGGCTGCATTTTAATATCACTTCCCCTCGCATTTTCCCATCAC	1960
Db	144324	AGTTACAGAGGGGCGCTGGGCTGCATTTTAATATCACTTCCCCTCGCATTTTCCCATCAC	144383
OY	1961	ATAGAAGACTTTGACCTGTGAAGCTGCACATCTGTTAATATCAATAATTTCCCAATTAAG	2017
Db	144384	ATAGAAGACTTTGACCTGTGAAGCTGCACATCTGTTAATATCAATAATTTCCCAATTAAG	144440
RESUTL	5	AF280088	1992 bp mRNA linear PRI 20-FEB-2001
LOCUS	AF280088	Homo sapiens I-selectin ligand sulfotransferase GST-3 mRNA,	
DEFINITION		complete cds.	
ACCESSION		AF280088	
VERSION		AF280088.1 GI:12060807	
KEYWORDS			
SOURCE			
ORGANISM		Homo sapiens.	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 1992)	
JOURNAL		Hemmerich,S., Lee,J.K., Bhakta,S., Bistrup,A., Ruddle,N.R. and	
MEDLINE		Rosen,S.D.	
PUBMED		Chromosomal localization and genomic organization for the	
REFERENCE		galactose/N-acetylgalactosamine/N-acetylglucosamine	
AUTHORS		6-O-sulfotransferase gene family	
TITLE		Glycobiology 11 (1), 75-87 (2001)	
JOURNAL		21096027	
MEDLINE		11181564	
PUBMED		2 (bases 1 to 1992)	
REFERENCE		Hemmerich,S., Bhakta,S., Lee,J.-K., Bistrup,A., Ruddle,N.R. and	
AUTHORS		Rosen,S.D.	
TITLE		Direct SubMISSION	
JOURNAL		Submitted (20-JUN-2000) Respiratory Diseases, Roche Bioscience,	
MEDLINE		3401 Hillview Ave., Palo Alto, CA 94304, USA	
PUBMED		location/Qualifiers	
FEATURES		. . .1992	
SOURCE		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="16"	
		/map="16q23.1-q23.2"	
		/cell_type="high endothelial"	
		/tissue_type="tonsil"	
		. . .1992	
		/gene="GST3"	
		122..1282	
		/note="similar to Homo sapiens sequence encoded by Genbank	
		accession Number AF131235"	
		/codon_start=1	
		/product="I-selectin ligand sulfotransferase GST-3"	
		/protein_id="AAG48246.1"	
		/db_xref="GI:12060808"	
		/translation="MLPRLKMILLFLVSONAILALFFRMYSHNITSLSMKRQPERMT	
		VLVLSWRSGSFVGQLFGQHPDVYLMEPAHWVMFTKOSTAMLMHAIVDLIRAVV	
gene			
CDs			

LCDSVPDPMYMEPRROSSIFQWENSHALUSAPACDIIPDEITIPRACHLLCSOP
PEVEKACRSTSHVLYKREVEFNLOSLYPLKIDSLNHIHVLVDPAPVRSERIK
GDMIDSRITVMGQHEQKLKEDQYVQVQSOQLKLEYKTIQSLPKJQRYLLVRY
EDLAAAPVOTSRMYEVEFLPHLOTWVNIQSLRGKMGHATHTNADLANYSQAM
RWSLPEYVSLKACGDANMLIGYRHVRSQEDRNLLDLSTWYVEQJH"
BASE COUNT 482 a 540 c 489 g 481 t
ORIGIN

Query Match 86.4%; Score 1756; DB 9; Length 1992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1856; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 160 CAGGCTTCCACTTCAGACATGCTACTGCTAAAAAATGAGCTCCCTGCTTCT 219
DB 100 CAGGCTTCCACTTCAGACATGCTACTGCTAAAAAATGAGCTCCCTGCTTCT 159
QY 220 GGTTCGCCAGATGGCCATCTGGCTCTATCTTCACATGTACAGCCACAACATCAGCTC 279
DB 160 GGTTCGCCAGATGGCCATCTGGCTCTATCTTCACATGTACAGCCACAACATCAGCTC 219
QY 280 CCGTCTATGAGGAGCAGACCCGAGCAGATGACGTGCTGCTGCTCTCTCTGCGCTC 339
DB 220 CCGTCTATGAGGAGCAGACCCGAGCAGATGACGTGCTGCTGCTCTCTCTGCGCTC 279
QY 340 TGGCTCTCTCTTGTGGGCGACCTTTTGGGCGAGCACCAGATGTTTCTACCTGATGA 399
DB 280 TGGCTCTCTCTTGTGGGCGACCTTTTGGGCGAGCACCAGATGTTTCTACCTGATGA 339
QY 400 GCGCCCTTGGCAGCTGTGATACCTTCAGAGAGCAGCCGCTTGATGTGACATGCG 459
DB 340 GCGCCCTTGGCAGCTGTGATACCTTCAGAGAGCAGCCGCTTGATGTGACATGCG 399
QY 460 TGTGGGGATGTGATACGGGCGCTCTTGTGCGACATGACGCTTTGATGCTTACAT 519
DB 400 TGTGGGGATGTGATACGGGCGCTCTTGTGCGACATGACGCTTTGATGCTTACAT 459
QY 520 GGAACCTGTGTCGCCGAGACAGCTCCAGCCTTTTCAGTGGGAGAAAGCCGGGCTGTG 579
DB 460 GGAACCTGTGTCGCCGAGACAGCTCCAGCCTTTTCAGTGGGAGAAAGCCGGGCTGTG 519
QY 580 TTCTGACACCTGCTGTGACATCATCCCAAGATGAAATCATCCCGGGGCTCACTGAG 639
DB 520 TTCTGACACCTGCTGTGACATCATCCCAAGATGAAATCATCCCGGGGCTCACTGAG 579
QY 640 GCTCTGTGACATCAACAGCCCTTTAGGTGTGTGAGAGAGGCTGCGCTCTTACAGCA 699
DB 580 GCTCTGTGACATCAACAGCCCTTTAGGTGTGTGAGAGAGGCTGCGCTCTTACAGCA 639
QY 700 CGTGTGTCAGAGGAGGCGCTTCTTCACTGCACTGACCTCCGCTGCTGGAAGA 759
DB 640 CGTGTGTCAGAGGAGGCGCTTCTTCACTGCACTGACCTCCGCTGCTGGAAGA 699
QY 760 CCGCTCCCTCAACCTGATATGTCACACTGATCCGGAGCCCGGCGGTGTCCTTC 819
DB 700 CCGCTCCCTCAACCTGATATGTCACACTGATCCGGAGCCCGGCGGTGTCCTTC 759
QY 820 CCGAAGAGCAGAAAGGAGATCTATATGACAGTGTGATGTGATGGGCGACATGA 879
DB 760 CCGAAGAGCAGAAAGGAGATCTATATGACAGTGTGATGTGATGGGCGACATGA 819
QY 880 GCAAGAACTCAAGAGAGGAGCAACCTACTATGTGATGAGAGTCACTTCCCAAGCA 939
DB 820 GCAAGAACTCAAGAGAGGAGCAACCTACTATGTGATGAGAGTCACTTCCCAAGCA 879
QY 940 GCTGAGATCTACAGACATCCATCTTGTGCCAAGGCGCTGCAAGAGCTTACTGCT 999
DB 880 GCTGAGATCTACAGACATCCATCTTGTGCCAAGGCGCTGCAAGAGCTTACTGCT 939
QY 1000 TGTGGCATGTAGAGCTGTGCTGAGCCCTGTGAGCCAGATTCGCCAATGTATGATT 1059
DB 940 TGTGGCATGTAGAGCTGTGCTGAGCCCTGTGAGCCAGATTCGCCAATGTATGATT 999
QY 1060 CGTGGGATGTGAATCTTGTGCCCATCTTTCAGACCTGGGTGATTAACATCACCAGCA 1119

DB 1000 CGTGGGATGTGAATCTTGTGCCCATCTTTCAGACCTGGGTGATTAACATCACCAGCA 1059
QY 1120 GGGCATGTGTGACACAGCTTTTCCACAGAAATGCCAGGATGCCCTTAATGCTCCAGG 1179
DB 1060 GGGCATGTGTGACACAGCTTTTCCACAGAAATGCCAGGATGCCCTTAATGCTCCAGG 1119
QY 1180 TTGGCGTGTGCTTTTGCCCTATGAAAAGTTTTCGACTTCAGAAAAGCCTGTGCGATGC 1239
DB 1120 TTGGCGTGTGCTTTTGCCCTATGAAAAGTTTTCGACTTCAGAAAAGCCTGTGCGATGC 1179
QY 1240 CATGATTTTGTGGCTTACCCGCAAGCTGATGATGACAAACAGAGAAACCTGTGCT 1299
DB 1180 CATGATTTTGTGGCTTACCCGCAAGCTGATGATGACAAACAGAGAAACCTGTGCT 1239
QY 1300 GGATCTTCTGTCTACTGACTGTGCTGAGCAATCCACTAAGAGGTTGAGAAGGCT 1359
DB 1240 GGATCTTCTGTCTACTGACTGTGCTGAGCAATCCACTAAGAGGTTGAGAAGGCT 1299
QY 1360 TGTGTCCACCTGTGTGACGCTCACTGCTTCTGTGATGCTTCAAGCCTTGCTACA 1419
DB 1300 TGTGTCCACCTGTGTGACGCTCACTGCTTCTGTGATGCTTCTGAGCCTTGCTACA 1359
QY 1420 TCTGTGACCTTTACTACATGCTGTGGGTATTCACACTGATGTGAGTGTGCTACAG 1479
DB 1360 TCTGTGACCTTTACTACATGCTGTGGGTATTCACACTGATGTGAGTGTGCTACAG 1419
QY 1480 TGTCTAACAGAGAGACTTTTGTGTCCATGCTTGTGTACAAAAACAGACTGGGAACT 1539
DB 1420 TGTCTAACAGAGAGACTTTTGTGTCCATGCTTGTGTACAAAAACAGACTGGGAACT 1479
QY 1540 TATGTGACAGACATCCACAGAGGATGAAAGGCTTGTGCTTCTTCTTGTGATCT 1599
DB 1480 TATGTGACAGACATCCACAGAGGATGAAAGGCTTGTGCTTCTTCTTGTGATCT 1539
QY 1600 TCTGTGTGCGAGACTTCAGAGACTTTGTGGCCTGAGAGGCTTATTAAGCAGCAGACT 1659
DB 1540 TCTGTGTGCGAGACTTCAGAGACTTTGTGGCCTGAGAGGCTTATTAAGCAGCAGACT 1599
QY 1660 ATCAGTGAATGTATTCATTAACCTCCCTGTCCACATCTTCCCAATGGGGAATGATCT 1719
DB 1600 ATCAGTGAATGTATTCATTAACCTCCCTGTCCACATCTTCCCAATGGGGAATGATCT 1659
QY 1720 TTCAACAAAGAGCTCACAGCAATTTTCCACAGATCCAAATTTGAGCCCTTGAATTC 1779
DB 1660 TTCAACAAAGAGCTCACAGCAATTTTCCACAGATCCAAATTTGAGCCCTTGAATTC 1719
QY 1780 CCAATGTGATTAAGAGAGAAAGTGGAAACAAGTTGATGCTTACTTATGAGCTTGACA 1839
DB 1720 CCAATGTGATTAAGAGAGAAAGTGGAAACAAGTTGATGCTTACTTATGAGCTTGACA 1779
QY 1840 TCACAGCTATGCTATACAGAAATATGAAACAAATCTGACACAAAAGCAAGCTCT 1899
DB 1780 TCACAGCTATGCTATACAGAAATATGAAACAAATCTGACACAAAAGCAAGCTCT 1839
QY 1900 AAGTTTCAAGGGTGTGCTGCTGATTTGAATATCACTTCCCTGATTTTCCATCA 1959
DB 1840 AAGTTTCAAGGGTGTGCTGCTGATTTGAATATCACTTCCCTGATTTTCCATCA 1899
QY 1960 CATAGAGACTTTTACCTGTGAAGCTCCATCTGATTAATCTAAATTTCCCAATAG 2017
DB 1900 CATAGAGACTTTTACCTGTGAAGCTCCATCTGATTAATCTAAATTTCCCAATAG 1957

RESULT 6
AF149783
LOCUS AF149783 1333 bp mRNA linear PRI 02-JUL-2001
DEFINITION Homo sapiens L-selectin ligand sulfotransferase mRNA, complete cds.
ACCESSION AF149783
VERSION AF149783.1 GI:13897503
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens

QY	161	AAATGCTTCACATTGACGACATGCTACGCGCTAAAAAATGAAGCTCCTGCTTCTG	220
Db	72	AAAGTCTTCACATTGACGACATGCTACTGCTAAAAAATGAAGCTCCTGCTTCTG	131
QY	221	GTTTCCAGATGGGCATCTTGAGCTCTATATCTTCCACATGTAAGCCACACATCAGCTCC	280
Db	132	GTTTCCAGATGGGCATCTTGAGCTCTATATCTTCCACATGTAAGCCACACATCAGCTCC	151
QY	281	CTGACTATGAAGGACAGCCCGAGAGGACATGACAGTGCAGTGTCTGCTTCCAGGAGCTC	340
Db	192	CTGACTATGAAGGACAGCCCGAGAGGACATGACAGTGTCTGCTTCCAGGAGCTC	251
QY	341	GCGCTTCTTTTGTGGGGCAGCTTTTGGGCGAGCCACAGATGTTTCTCACTGATGAG	400
Db	252	GCGCTTCTTTTGTGGGGCAGCTTTTGGGCGAGCCACAGATGTTTCTCACTGATGAG	311
QY	401	CCCGCCTGGCAGCATGTGGATGACCTTCAAGCAGAGACACCGCGCTGATGCTGCACATGCT	460
Db	312	CCCGCCTGGCAGCATGTGGATGACCTTCAAGCAGAGACACCGCGCTGATGCTGCACATGCT	371
QY	461	GTCGGGATGTGATACGGGCGCGCTCTCTGTGTCGACATGAGCGCTTGAATGCTAATG	520
Db	372	GTCGGGATGTGATACGGGCGCGCTCTCTGTGTCGACATGAGCGCTTGAATGCTAATG	431
QY	521	GAACCTGTCCTCCGGAGACAGTCACGCGCTTTCAGTGGGGAAGACAGCGGGCGCTGTGT	580
Db	432	GAACCTGTCCTCCGGAGACAGTCACGCGCTTTCAGTGGGGAAGACAGCGGGCGCTGTGT	491
QY	581	TCTGACACTCCCTGTACATCATCCACCAAGATGAATCATCCCGGGGCTCACTGAGG	640
Db	492	TCTGACACTCCCTGTACATCATCCACCAAGATGAATCATCCCGGGGCTCACTGAGG	551
QY	641	CTCCTGTGACATCAACAGCCCTTTGAGGTGTGGAAAGSCCTCGCGCTCTACAGCCAC	700

Db	552	CTCCGTCGACATCAACAGCCCTTTGAGGTGGTGGAGAAAGGCGCTGCGGCTCCTACAGCCAC	611
Qy	701	GTGGGTCGCAAGAGAGTGGCGCTTCTTCAACCTGGAGTGCCTCTACCCGCTCTGAAGAC	760
Db	612	GTGGGTCGCAAGAGAGTGGCGCTTCTTCAACCTGCAATGCTCTTACCCGCTCTGAAGAC	671
Qy	761	CCCTCCCTCAACCTGCATATCGTGCACCTGGTCCGGGACCCCGGGCGGTTCGTTCC	820
Db	672	CCCTCCCTCAACCTGCATATCGTGCACCTGGTCCGGGACCCCGGGCGGTTCGTTCC	731
Qy	821	CGAGAACCCCAAAAGGAGATCTCATGTATACATGTCATTTGATGATGGGACAGATAG	880
Db	732	CGAGAACCCCAAAAGGAGATCTCATGTATACATGTCATTTGATGATGGGACAGATAG	791
Qy	881	CAGAAACTCAAGAAAGAGAGACCAACCCCTACTATGTGAATGACAGTCACTGTGCCAAAGCCAG	940
Db	792	CAAAACCTCAAGAAAGAGAGACCAACCCCTACTATGTGAATGACAGTCACTGTGCCAAAGCCAG	851
Qy	941	CTGAGATCTCAAGAACCATCCAGTCTTGGCCCAAGGCCCTGACAGAAAGCTTACTGCTT	1000
Db	852	CTGAGATCTCAAGAACCATCCAGTCTTGGCCCAAGGCCCTGACAGAAAGCTTACTGCTT	911
Qy	1001	GTGGCCTTGAAGGACCTGGCTCGAGCCCTGCGGCCACAGCTCCCAATGATGAATTC	1060
Db	912	GTGGCCTTGAAGGACCTGGCTCGAGCCCTGCGGCCACAGCTCCCAATGATGAATTC	971
Qy	1061	GTGGGATTGAATTCCTGCCCATCTTCAGACCTGGGTGATTAACATCACCCGAGGCAAG	1120
Db	972	GTGGGATTGAATTCCTGCCCATCTTCAGACCTGGGTGATTAACATCACCCGAGGCAAG	1031
Qy	1121	GGCATGGGTGACACAGCTTTTCCACCAAAATGCGAGAGATGGCCCTTAATGTCTCCAGGCT	1180
Db	1032	GGCATGGGTGACACAGCTTTTCCACCAAAATGCGAGAGATGGCCCTTAATGTCTCCAGGCT	1091
Qy	1181	TGGCCCTGGTCTTGGCCCTATGAAGAAGTTTCTGACTTCAGAAAGCCCTGTGGCGATGCC	1240
Db	1092	TGGCCCTGGTCTTGGCCCTATGAAGAAGTTTCTGACTTCAGAAAGCCCTGTGGCGATGCC	1151
Qy	1241	ATGAATTTGCTGGGCTACCGGCACGTCAATGTGAACAAGAACAGAAACCTGTTGCTG	1300
Db	1152	ATGAATTTGCTGGGCTACCGGCACGTCAATGTGAACAAGAACAGAAACCTGTTGCTG	1211
Qy	1301	GATCTCTGTCTACTGAGACGTCCCTAGAGAAATCCACTAAAGAGGTTGAAGAAGCTTT	1360
Db	1212	GATCTCTGTCTACTGAGACGTCCCTAGAGAAATCCACTAAAGAGGTTGAAGAAGCTTT	1271
Qy	1361	GCTGCCACCTGGTGTGACGCTCAAGTCACTTCTCTGAATGCTCTGAGCCTTGCTACAT	1420
Db	1272	GCTGCCACCTGGTGTGACGCTCAAGTCACTTCTCTGAATGCTCTGAGCCTTGCTACAT	1331
Qy	1421	CT 1422	
Db	1332	CT 1333	
RESULT 7			
AX381256	LOCUS	517 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 194 from Patent WO0212280.		
AX381256	ACCESSION		
VERSION	AX381256.1 GI:19576075		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	Pyle, R.A., Xu, J. and Secrist, H.		
JOURNAL	Compositions and methods for the therapy and diagnosis of colon		
FEATURES	Patent: WO 0212280-A 1994 14-FEB-2002;		
	CORIXA CORPORATION (US)		
	Location/Qualifiers		

CJMDVEIDAIMQSRNLSAFENWATSRALCSPPACSAFPRGITSKQDCKTLCTROPES
LAREACRSYSHVLEKVRFFENIQVLYPLSLSPALNRIHVLRDPRAVLRSREACPT
LARDNGIVLGTNGKWEADPHRLIREVCRSHVRIAEATLKPDPPLDGRYRLVREED

BASE COUNT 239 a 510 c 446 g 267 t
ORIGIN
Query Match 2.9%; Score 59; DB 9; Length 1462;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 736
|||||
Db 719 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 777

RESULT 10
AX327330 1647 bp DNA linear PAT 07-JAN-2002
LOCUS Sequence 15 from Patent WO0179468.
DEFINITION AX327330
ACCESSION AX327330
VERSION AX327330.1 GI:18097876
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Policky, J.L., Hafalia, A., Burford, N., Ring, H.Z., Lal, P.,
Tribouley, C.M., Yao, M.G., Yue, H., Tang, Y.T., Patterson, C., Das, D.,
Sanjanwala, M.S., Gandhi, A.R., Reddy, R., Khan, F.A., Baughn, M.R.,
Ramkumar, J., Griffin, J.A. and Au-Yang, J.
TITLE Drug metabolizing enzymes
JOURNAL Patent: WO 0179468-A 15 25-OCT-2001;
Incyte Genomics, Inc. (US)

FEATURES
source
1..1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7472777CB1"
BASE COUNT 287 a 560 c 501 g 299 t
ORIGIN

Query Match 2.9%; Score 59; DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 736
|||||
Db 488 AGGCTGCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 546

RESULT 11
AF176838 2170 bp mRNA linear PRI 22-SEP-1999
LOCUS AF176838
DEFINITION Homo sapiens intestine N-acetylglucosamine 6-O-sulfotransferase
(I-G1CNAc-6-ST) mRNA, complete cds.
ACCESSION AF176838
VERSION AF176838.1 GI:5917705
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Lee, J.K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
TITLE Cloning and characterization of a mammalian
N-acetylglucosamine-6-sulfotransferase that is highly restricted to
intestinal tissue
JOURNAL Biochem. Biophys. Res. Commun. 263 (2), 543-549 (1999)
MEDLINE 99423499
PUBMED 10491328
REFERENCE 2 (bases 1 to 2170)
AUTHORS Lee, J.-K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
TITLE Direct Submission

JOURNAL Submitted (10-AUG-1999) Respiratory Diseases, Roche Bioscience,
3401 Hillview Avenue, Palo Alto, CA 94304, USA
FEATURES
source
1..2170
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q23.1"
/tissue_type="intestine"
1..2170
/gene="I-G1CNAc-6-ST"
344..1516
/gene="I-G1CNAc-6-ST"
/note="possibly involved in biosynthesis of mucosal
ligands for L-selectin"
/codon_start=1
/product="N-acetylglucosamine 6-O-sulfotransferase"
/protein_id="A0D56000.1"
/db_xref="GI:5917706"
/translation="MWLPFRSSKTIVVLLAQTCLLFITSPRPSSPAGEDRVH
LVLSMRGSSFLIGLFSQHPDVFYLMPEAMHWTLTSGSAATLHMAVDIMRSIFL
CDMDVFDYMPQSNLSFTFMATSRALCSRPACSAFPRGTSIKDYCKRLCTRQPS
LARECRSHVLEKEVFNLQVLPDLSPLNLKLVLRPRAVLRSRAGFI
LARGNIVLGTNGKWEADPHLRILRVCRSHVRIAAATLKLPPPLRGYRIWFED
LAREPLAEIRALVAFGLTLPQLEAMHNTTHSGSGIKPIEAFRTSSRNARVSCAM
RHAPFTKILRVQECAGALQILGRVYSADQORDLTLVLRPDPHFSWASPD"

BASE COUNT 449 a 682 c 612 g 427 t
ORIGIN

Query Match 2.9%; Score 59; DB 9; Length 2170;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 736
|||||
Db 834 AGGCTGCCGCTCTACAGCCAGCGTGTGCTCAAGAGGCGCTTCTTCACCTGCAG 892

RESULT 12
AF219990 2544 bp mRNA linear PRI 26-OCT-2000
LOCUS AF219990
DEFINITION Homo sapiens corneal N-acetylglucosamine-6-O-sulfotransferase
(CHST6) mRNA, complete cds.
ACCESSION AF219990
VERSION AF219990.1 GI:11023145
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Akama, T.O., Nishida, K., Nakayama, J., Watanabe, H., Fujiwara, T.,
Nakamura, T., Dote, A., Kawasaki, S., Inoue, Y., Maeda, N., Yamamoto, S.,
Ozaki, K., Kinoshita, S., Shimomura, Y., Tanigami, A. and Fukuda, M.N.
TITLE Macular corneal dystrophy type I and type II are caused by distinct
mutations in a new sulphotransferase gene
JOURNAL Nat. Genet. 26 (2), 237-241 (2000)
MEDLINE 20472350
PUBMED 11017086
REFERENCE 2 (bases 1 to 2544)
AUTHORS Akama, T.O. and Fukuda, M.N.
TITLE Direct Submission
JOURNAL Submitted (29-DEC-1999) Glycobiology Program, The Burnham
Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES
source
1..2544
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22"
1..2544
/gene="CHST6"
/note="responsible for macular corneal dystrophy"

CDS

693..1880
/gene="CHST6"
/function="Transfers sulfate to 6-OH of GlcNAc in keratan sulfate"
/note="C-GlcNAc6S7"
/codon_start=1
/product="corneal N-acetylglucosamine-6-O-sulfotransferase"
/protein_id="AA026325.1"
/db_xref="GI:11023146"
/translation="MDEPRVSSIVATALLAQTELLFLVSRPSSPAGEARVAVLVLSWRSSSSVQGLFNOHPDYFLMEAPAMVTTLSQSNATLHMAVRDVRSEFLCDDVDYDALPWRNLSDLEQMAVSRALCPSPGAPPGATISSEAVCKPSPGSEFLAREACRSYSHVLEKEVRENFNOVYPLISDPALNRLVHLVRDRAVLRSEQKALARDNGIVGTGTWEADPCGLRVREYCRSRVRAENATLKPPPLIGRRLVFEEDLAREPLAEIRALYATGSLTLPQLEWVHNTHSGPGARRAEKTSRNALNVSAWRHALPPAKIRRVQELCAGALQLGLGRPYVSEDEQRNALDLVLPGLNGFTWASSTASHPRN"

BASE COUNT 460 a 799 c 733 g 552 t

ORIGIN

Query Match 2.9% Score 59; DB 9; Length 2544;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGCTCTACACGCCAGCTGTCCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
|||||
DB 1180 AGGCGTGGCGCTCTACACGCCAGCTGTCCTCAAGAGAGTGGCTTCTTCAACCTGCAG 1238

RESULT 13
AF246718 3278 bp mRNA linear PRI 31-OCT-2000
LOCUS Homo sapiens intestinal GlcNAc-6-sulfotransferase (CHST5) mRNA,
DEFINITION complete cds, alternatively spliced.
ACCESSION AF246718
VERSION AF246718.1 GI:11055254
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3278)
AUTHORS Akama,T.O., Nishida,K., Nakayama,J., Watanabe,H., Ozaki,K., Nakamura,T., Dote,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S., Fujiwara,T., Thonar,E.J., Shimomura,Y., Kinoshita,S., Tanigami,A. and Fukuda,M.N.
TITLE Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulphotransferase gene
JOURNAL Nat. Genet. 26 (2), 237-241 (2000)
MEDLINE 20472330
PUBMED 11017086
REFERENCE 2 (bases 1 to 3278)
AUTHORS Akama,T.O., Nishida,K., Nakayama,J., Watanabe,H., Ozaki,K., Nakamura,T., Dote,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S., Fujiwara,T., Thonar,E.J., Shimomura,Y., Kinoshita,S., Tanigami,A. and Fukuda,M.N.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2000) Glycobiology Program, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES
source 1..3278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22: between D16S512 and D16S515"
1..3278
/gene="CHST5"
1..1396
/note="alternatively spliced"
1397..2632
/gene="CHST5"

/note="I-GlcNAc6ST; carbohydrate sulfoltransferase 5; GlcNAc6ST expressed in small intestine and colon; similar to N-acetylglucosamine 6-O-sulfotransferase encoded by GenBank Accession Number AF176838; alternatively spliced"

BASE COUNT 686 a 988 c 866 g 738 t

ORIGIN

Query Match 2.9% Score 59; DB 9; Length 3278;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGCTCTACACGCCAGCTGTCCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
|||||
DB 1950 AGGCGTGGCGCTCTACACGCCAGCTGTCCTCAAGAGAGTGGCTTCTTCAACCTGCAG 2008

RESULT 14
AF280086 3786 bp mRNA linear PRI 20-FEB-2001
LOCUS Homo sapiens N-acetylglucosamine 6-O-sulfotransferase GST-4beta
DEFINITION mRNA, complete cds.
ACCESSION AF280086
VERSION AF280086.1 GI:12060803
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3786)
AUTHORS Hemmerich,S., Lee,J.K., Bhakta,S., Bistrup,A., Ruddle,N.R. and Rosen,S.D.
TITLE Chromosomal localization and genomic organization for the galactose/N-acetylgalactosamine/N-acetylglucosamine 6-O-sulfotransferase gene family
JOURNAL Glycobiology 11 (1), 75-87 (2001)
MEDLINE 21096027
PUBMED 11181564
REFERENCE 2 (bases 1 to 3786)
AUTHORS Hemmerich,S., Bhakta,S., Lee,J.K., Bistrup,A., Ruddle,N.R. and Rosen,S.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2000) Respiratory Diseases, Roche Bioscience, 3401 Hillview Ave., Palo Alto, CA 94304, USA
FEATURES
source 1..3786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q23.1"
1..3786
/gene="GST4beta"
/note="clusters with the intestinal N-acetylglucosamine 6-O-sulfotransferase (GST4alpha) gene in GenBank Accession Numbers AF176838 and AF176839"
189..1376
/gene="GST4beta"
/note="similar to N-acetylglucosamine 6-O-sulfotransferase GST-4alpha"
/product="N-acetylglucosamine 6-O-sulfotransferase GST-4beta"

FEATURES	source
polya_signal	3106. 3111 /gene="Gst4beta" /note="alternative signal" 3734. 3739 /gene="Gst4beta"
BASE COUNT	848 a 1030 c 1104 g 804 t
ORIGIN	
Query Match	2.9%; Score 59; DB 9; Length 3786;
Best Local Similarity	100.0%; Pred. No. 1,1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
CY 678	AGGCGTGGCGGCTCAGACGCGAGGCGTGCACAGAGAGGCGTGCCTTCACCTGCAG 736
Db 676	AGGCGTGGCGGCTCAGACGCGAGGCGTGCACAGAGGCGTGCCTTCACCTGCAG 734
RESULT 15	
LOCUS	AF219991 71503 bp DNA linear PRI 26-OCT-2000
DEFINITION	Homo sapiens intestinal N-acetylglucosamine-6-O-sulfotransferase (CHST5) and corneal N-acetylglucosamine-6-O-sulfotransferase (CHST5) genes, complete cds.
ACCESSION	AF219991
VERSION	AF219991.1 GI:11023147
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 71503) Akama,T.O., Nishida,K., Nakayama,J., Watanabe,H., Fujiwara,T., Nakamura,T., Dotsa,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S., Ozaki,K., Kinoshita,S., Shimomura,Y., Tanigami,A. and Fukuda,M.N.
TITLE	Mucular corneal dystrophy type I and type II are caused by distinct mutations in a new sulfotransferase gene
JOURNAL	Nat. Genet. 26 (2), 237-241 (2000)
MEDLINE	20472330
PUBMED	11017086
REFERENCE	2 (bases 1 to 71503) Akama,T.O. and Fukuda,M.N.
AUTHORS	Direct Submission
TITLE	Submitted (29-DEC-1999) Glycobiology Program, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
JOURNAL	Location/Qualifiers
FEATURES	1. 71503 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q22" /clone="CITB-483K2" 358. 461 /rpt_family="L2" /rpt_type-dispersed complement(462..771) /rpt_family="Alu" /rpt_type-dispersed 772. 945 /rpt_family="L2" /rpt_type-dispersed 1169..1338 /rpt_family="Alu" /rpt_type-dispersed 15977..16098 /rpt_family="MIR"

```

repeat_region /rpt_type=dispersed complement(17289..17433)
/rpt_family="MIR"
repeat_region /rpt_type=dispersed complement(18001..18310)
/rpt_family="Alu"
repeat_region /rpt_type=dispersed complement(18420..18591)
/rpt_family="MIR"
repeat_region /rpt_type=dispersed 18732..18767
/rpt_family="Simple-repeat"
/rpt_type=tandem 18967..20202
gene="CHST5"
note="I-GlcNAc6ST; carbohydrate sulfotransferase 5; expressed in the small intestine and colon"
/codon_start=1
/product="intestinal N-acetylglucosamine-6-O-sulfotransferase"
/protein_id="AAG26326.1"
/db_xref="GI:11023148"
/translation="MGMRARVRYVAHSRRPPAARMMLPRFSSTKVIVLLAOTCIL
LEITSRPPSSPAGDEPRVHYLVISRRSSSPGQLESCHPDVFLMEPRWHTL
SOGSNATIHMAVRDLMRSTFLCDMDVFAIMPQSRNLSAFTNMTSRGLCPACSAF
PRGTSKQDVCKTCTQRFPSLAKACSHVVAKEVREFNLQVLPGLSDPLNR
IVHLVRDRAVLRSREAGPILANDNGIVLTNGKWEADPHRLIRVCRSHYRIE
AALIKPPFLRGRLVRFEDLAREPLAFLYATGLITLPDLEAMINHTGSGI
GKPIEAFSTSRNARNVQAMRHALPFTKILRVQEVCAQLLLGVRVVSADQQRDL
TLDLVPRGDPHFWASPD"
repeat_region 20504..20806
/rpt_family="Alu"
/rpt_type=dispersed 20869..20961
/rpt_family="L1"
/rpt_type=dispersed 20962..21200
/rpt_family="MER4-group"
/rpt_type=dispersed 21325..21429
/rpt_family="MER4-group"
/rpt_type=dispersed 21430..21722
/rpt_family="Alu"
/rpt_type=dispersed 21725..22007
/rpt_family="Alu"
/rpt_type=dispersed 22008..22040
/rpt_family="Simple-repeat"
/rpt_type=tandem 22041..22313
repeat_region /rpt_family="MER4-group"
/rpt_type=dispersed 22325..22593
/rpt_family="Alu"
/rpt_type=dispersed 22594..22949
/rpt_family="MER4-group"
/rpt_type=dispersed 23116..23221
/rpt_family="Simple-repeat"
/rpt_type=tandem 23344..23400
/rpt_family="MALR"
/rpt_type=dispersed 23433..23456
repeat_region /rpt_family="Simple-repeat"
/rpt_type=tandem complement(23457..23720)
/rpt_family="Alu"
/rpt_type=dispersed complement(23789..24140)

```

```

/rpt_family="MER21-group"
/rpt_type=dispersed 24141..24287
/rpt_family="MALR"
/rpt_type=dispersed complement(24308..24420)
/rpt_family="MALR"
/rpt_type=dispersed complement(24467..24528)
/rpt_family="MER1-type"

Query Match 2.9%; Score 59; DB 9; Length 71503;
Best Local Similarity 100.0%; Pred. No. 1.2e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

cy 678 AGCGCTGCGCTCCACACGACGCGTGCACAGAGAGTGCCTTCTCAACCTGCAG 736
db 19520 AGCGCTGCGCTCCACACGACGCGTGCACAGAGAGTGCCTTCTCAACCTGCAG 19578

RESULT 16
AC009163/c 157337 bp DNA linear PRI 02-NOV-2001
LOCUS AC009163
DEFINITION Homo sapiens chromosome 16 clone RP11-77K12, complete sequence.
ACCESSION AC009163
VERSION AC009163.5 GI:16596526
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE DOE Joint Genome Institute.
JOURNAL Unpublished
2 (bases 1 to 157337)
AUTHORS DOE Joint Genome Institute.
TITLE DOE Joint Genome Institute.
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 157337)
AUTHORS DOE Joint Genome Institute.
TITLE DOE Joint Genome Institute.
JOURNAL Submitted (02-NOV-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Nov 2, 2001 this sequence version replaced gi:17689976.
COMMENT
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source Location/Qualifiers
1..157337
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-77K12"

BASE COUNT 45295 a 36632 c 36308 g 39102 t

```

ORIGIN

Query Match 2.9%; Score 59; DB 9; Length 157337;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGCTCTCTACAGCAGCTGCTCAAGAGAGCGCTTCTTCAACCTGCAG 736
|||||
Db 144095 AGGCTGCGCTCTCTACAGCAGCTGCTCAAGAGAGCGCTTCTTCAACCTGCAG 144037
|||||

RESULT 17

LOCUS AC025287 194832 bp DNA 1linear PRI 26-JAN-2002
AC025287
DEFINITION Homo sapiens chromosome 16 clone RP11-490B18, complete sequence.
ACCESSION AC025287 GI:18376863
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
JOURNAL
REFERENCE
AUTHORS
DIRECT SUBMISSION
2 (bases 1 to 194832)
DOE Joint Genome Institute.
TITLE
JOURNAL
REFERENCE
AUTHORS
DIRECT SUBMISSION
Submitted (08-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
JOURNAL
REFERENCE
AUTHORS
DIRECT SUBMISSION
Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
JOURNAL
REFERENCE
AUTHORS
DIRECT SUBMISSION
Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 26, 2002 this sequence version replaced gi:17976465.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.6.
Location/Qualifiers
1. 194832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-490B18"

BASE COUNT 53784 a 43240 c 43815 g 53993 t

ORIGIN

Query Match 2.9%; Score 59; DB 9; Length 194832;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGCTCTCTACAGCAGCTGCTCAAGAGAGCGCTTCTTCAACCTGCAG 736
|||||
Db 188038 AGGCTGCGCTCTCTACAGCAGCTGCTCAAGAGAGCGCTTCTTCAACCTGCAG 188096
|||||

RESULT 18

LOCUS AC009105 208185 bp DNA 1linear HTG 25-APR-2001
AC009105
DEFINITION Homo sapiens chromosome 16 clone RP11-455E15, WORKING DRAFT
SEQUENCE 17 unordered pieces.
ACCESSION AC009105
VERSION AC009105.7 GI:13786304

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVE1N.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 208185)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 208185)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 25, 2001 this sequence version replaced gi:7689928.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 0
Center clone name: RP11-455E15

Summary Statistics
Consensus quality: 180040 bases at least Q40
Consensus quality: 192095 bases at least Q30
Consensus quality: 195726 bases at least Q20
Estimated insert size: 195850; agarose-fp estimation
Estimated insert size: 206585; sum-of-contigs estimation
Quality coverage: 6.57 in Q20 bases; agarose-fp estimation
Quality coverage: 6.23 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2359: contig of 1197 bp in length
* 2359 2459: gap of unknown length
* 2459 2460 3653: contig of 1194 bp in length
* 3653 3654 3753: gap of unknown length
* 3753 3754 5640: contig of 1887 bp in length
* 5640 5740: gap of unknown length
* 5740 5741 7370: contig of 1630 bp in length
* 7370 7371 7470: gap of unknown length
* 7470 7471 8584: contig of 1114 bp in length
* 8584 8684: gap of unknown length
* 8684 8685 10014: contig of 1330 bp in length
* 10014 10115: gap of unknown length
* 10115 12451: contig of 2337 bp in length
* 12451 12452 12551: gap of unknown length
* 12551 28715: contig of 16164 bp in length
* 28715 28716 28815: gap of unknown length
* 28815 28816 43240: contig of 14325 bp in length
* 43240 43241 53865: contig of 10625 bp in length
* 53865 53866 53965: gap of unknown length
* 53965 53966 72091: contig of 18126 bp in length
* 72091 72092 72192: gap of unknown length
* 72192 72193 87356: contig of 15164 bp in length
* 87356 87357 87455: gap of unknown length
* 87455 87456 109461: contig of 22006 bp in length
* 109461 109462 131379: contig of 21818 bp in length
* 131379 131380 131479: gap of unknown length
* 131479 131480 157556: contig of 26077 bp in length
* 157556 157557 157558: gap of unknown length
* 157558 208185: contig of 50529 bp in length.

FEATURES
source Location/Qualifiers
1..208185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-455E15"
/clone_lib="RPC1 human BAC library 11"
BASE COUNT 58183 a 45373 c 45592 g 57204 t 1833 others
ORIGIN

Query Match 2.9%; Score 59; DB 2; Length 208185;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTCCGCTCTACAGCCAGTGGTCTCAGAGAGTGGCTTTCTCACTGCAG 736
|||||
Db 107982 AGGCGTCCGCTCTACAGCCAGTGGTCTCAGAGAGTGGCTTTCTCACTGCAG 108040
|||||

RESULT 19
AR203340 37 bp DNA linear PAT 20-JUN-2002
LOCUS AR203340
DEFINITION Sequence 7 from patent US 6365365.
ACCESSION AR203340
VERSION AR203340.1 GI:21499704
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 37)
AUTHORS Bistrup, A., Rosen, S.D., Tangemann, K. and Hemmerich, S.
TITLE Method of determining whether an agent modulates glycosyl
JOURNAL sulfotransferase-3
Patent: US 6365365-A 7 02-APR-2002;
FEATURES Location/Qualifiers
source 1..37
/organism="unknown"
BASE COUNT 14 a 9 c 8 g 6 t
ORIGIN

Query Match 1.8%; Score 37; DB 6; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTAAGAAGAGAGACCAACCTACTATGTGATGC 920
|||||
Db 1 AAACCTAAGAAGAGAGACCAACCTACTATGTGATGC 37
|||||

RESULT 20
AF131236 1926 bp DNA linear ROD 10-AUG-1999
LOCUS AF131236
DEFINITION Mus musculus N-acetylglucosamine 6-O-sulfotransferase gene,
complete cds.
ACCESSION AF131236
VERSION AF131236.1 GI:4927115
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1926)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1926)
JOURNAL Bistrup, A., Bhakta, S., Lee, J.K., Belov, Y.Y., Gunn, M.D., Zuo, F.R.,
Huang, C.C., Kannagi, R., Rosen, S.D. and Hemmerich, S.
Sulfotransferases of two specificities function in the
reconstitution of high endothelial cell ligands for L-selectin
J. Cell Biol. 145 (4), 899-910 (1999)
MEDLINE 99264336
PUBMED 10330415
REFERENCE 2 (bases 1 to 1926)
AUTHORS Bistrup, A., Tangemann, K., Bhakta, S., Lee, J.-K., Belov, Y.Y.,
Gunn, M.D., Zuo, F.-R., Huang, C.-C., Kannagi, R., Rosen, S.D. and
Hemmerich, S.

TITLE Direct Submission
JOURNML Submitted (25-FEB-1999) Respiratory Diseases, Roche Bioscience,
3401 Hillview Avenue, Palo Alto, CA 94304, USA
FEATURES
source Location/Qualifiers
1..1926
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/cell_type="high endothelial"
/tissue_type="tonsil"
<405..>1571
/product="N-acetylglucosamine 6-O-sulfotransferase"
405..1571
/function="biosynthesis of endothelial ligands for
L-selectin"
/note="HEC-G1CNC-6-ST"
/codon_start=1
/product="N-acetylglucosamine 6-O-sulfotransferase"
/protein_id="P4033016.1"
/db_xref="GI:4927116"
/translation="MMLKKGRLLMFAGSQVIVVAFIHMSVERHLSORESRPVAV
LVLSMRSGSSFVGOLEFGHPDVEFLKEPAMWPFSTAWKLRHAYADLRVEL
CDMSVPDAPVNNPGRRKQSSLFQWFOSSALCSAPVCPDPHAKHSRKHCKLGGQPF
DMYKACRSHGFPVLAKEVRFSLDQALPLITDPSLNIHVHLYRDPRAVFRSEHTI
ELWDSHIVIGQHEJTEKEDQPTAKTKCSQVDIVKVIQTLPELQOQVFLFVE
DLVAPLAQTTRLYKRVGIDFLPHLQVWVYVYRGMQGHAFHTRNRLNYSQAMR
WSLPEYEVSOLODPCGEMDLGLYQVRSQOEQGNLSLDLSSSHLIGQVREG"

BASE COUNT 426 a 520 c 484 g 496 t
ORIGIN

Query Match 1.3%; Score 27; DB 10; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGTCTTGATGCTACATG 520
|||||
Db 714 GACATGAGCGTCTTGATGCTACATG 740
|||||

RESULT 21
AF109155 2201 bp mRNA linear ROD 27-JUL-1999
LOCUS AF109155
DEFINITION Mus musculus L-selectin ligand sulfotransferase mRNA, complete cds.
ACCESSION AF109155
VERSION AF109155.1 GI:5596405
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2201)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 2201)
JOURNAL Hiraoka, N., Petryniak, B., Nakayama, J., Tsuboi, S., Suzuki, M.,
Yeh, J.C., Izawa, D., Tanaka, T., Miyasaka, M., Lowe, J.B. and Fukuda, M.
A novel, high endothelial venule-specific sulfotransferase
expresses 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed
by CD34
Immunity 11 (1), 79-89 (1999)
MEDLINE 99361934
PUBMED 10435581
REFERENCE 2 (bases 1 to 2201)
AUTHORS Hiraoka, N. and Fukuda, M.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-1998) Glycobiology, The Burnham Institute, 10901
North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES
source Location/Qualifiers
1..2201
/organism="Mus musculus"
/db_xref="taxon:10090"
319..1485
/function="directs expression of 6-sulfo sialyl Lewis x,
an L-selectin ligand displayed by CD34"
/note="tssr"
/codon_start=1

```

/product="L-selectin ligand sulfotransferase"
/protein_id="A045579.1"
/db_xref="GI:5396406"
/translation="MMLKKGRLMFLGQVIVVAFIMSVYRHLSQREESRRPYVH
LVLSWRSGSSSEVGGQLFGQHPDYFLMEPAWHYMTFTSTAKLHMAYRDLRSYFL
CDMSVFDAYNNPGPRKSSLPQWESRALCSAPVCDFFPAHEISSPKHCKLLGGQDPF
DWEKACRSHGSEVVLKREVRPLSQALYPLTDPSLNLHVHLYRDPRAVRSREHPTI
ELVVDSHIVIGCHLEPIKEEDOPYAMKTIICKSVDIVKAIOTLPBALOQRYLFLAYE
DIYRAFLAQTTRILKFTVGIDFLPHLOTWYHNTYTRGKMGQHAFHTNARNALNVSQAMR
WSLPYERVSQLDAGCEAMDLLGYLQVRSQDEQGNLSDLLSSSHLGOVREG"

```

```

BASE COUNT      537 a      629 c      534 g      501 t
ORIGIN
Query Match      1.3%; Score 27; DB 10; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 494 GACATGAGCGTCTTTGATGCTTACATG 520
      |||
Db 628 GACATGAGCGTCTTTGATGCTTACATG 654

```

Search completed: February 22, 2003, 22:24:25
 Job time : 6247 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 18:56:30 ; Search time 444 Seconds

(without alignments)
10306.443 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032
Sequence: 1 ggcctcgagccagcagatgcct.....ataagaaaaaaaaaaaaa 2032

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 25

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2032	100.0	2032	20	AAZ20792	Human glycosyl sul
2	1802	88.7	1979	22	AAK94229	Human full-length
3	1615	79.5	2065	21	AAZ94211	Human transferrase
4	1211	59.6	1333	24	AA516947	Human L-selectin s
5	741	36.5	877	22	AAK91803	Human CDNA 5'-end
6	741	36.5	877	22	AAK93921	Human CDNA clone r
7	505	24.9	517	24	AAK54724	Human colon cancer
8	351	17.3	548	22	AAK92568	Human CDNA 3'-end
9	306	15.1	2988	21	AACT6156	Human ORF1711

10	286	14.1	389	24	AAK54794	Human colon Cancer
11	60	3.0	60	24	AAK38074	Human spliced tran
12	59	2.9	1647	24	AAAD24670	Human drug metabol
13	59	2.9	1694	22	AAAD02700	Human glycosyl sul
14	59	2.9	2044	22	AAAD02699	Human glycosyl sul
15	59	2.9	2170	22	AAAD02698	Human glycosyl sul
16	59	2.9	2344	24	AAK89506	Human corneal N-ac
17	59	2.9	48436	24	AAK89533	Human corneal N-ac
18	59	2.9	160552	22	AAAD02697	Human glycosyl sul
19	37	1.8	37	20	AAZ20798	PCR primer for gly
20	30	1.5	30	24	AAK70821	probe for human ge
21	27	1.3	1926	20	AAZ20793	Mouse glycosyl sul

ALIGNMENTS

RESULT 1
AAZ20792
ID AAZ20792 standard; DNA; 2032 BP.
XX
AC AAZ20792;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human glycosyl sulfotransferase-3 coding sequence.
XX
KW glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KW selectin binding interaction; inflammation; lymphocyte homing; human;
KW secondary lymph organ; ss.
XX
OS Homo sapiens.
XX
PN WO949018-A1.
XX
PD 30-SEP-1999.
XX
PE 26-FEB-1999; 95WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIV CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistup A, Rosen SD, Tangemann K, Hemmerich S;
XX WPI, 1999-580442/49.
DR P-PDB; AAY39918.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
XX polynucleotides -
XX
PS Claim 4; Fig 1; 59pp; English.
XX
CC This sequence encodes the human glycosyl sulfotransferase-3 (GST-3) of
CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC The sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions;
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.
XX
SQ Sequence 2032 BP; 468 A; 569 C; 490 G; 505 T; 0 other;
XX
Query Match 100.0%; Score 2032; DB 20; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCCTCGAGCCAGATGCTCCAGTCGCGGGAATGCTTCCTCATTTGCTCCAG 60
Db 1 GGCCTCGAGCCAGATGCTCCAGTCGCGGGAATGCTTCCTCATTTGCTCCAG 60

QY	1141	CGACAAAGATCCCGGGATATCCCTTAATGTCCTCCAGGGTTGGCGCTGTCTTTCCCTTA	1200
Db	1141	CCACACAAATCCCGGGATATCCCTTAATGTCCTCCAGGGTTGGCGCTGTCTTTCCCTTA	1200
QY	1201	TGAAAAGGTTCTCGACTTCAGAAAGCCTGTGGCGATGCCATGAATTTCCTGGGCTACCG	1260
Db	1201	TGAAAAGGTTCTCGACTTCAGAAAGCCTGTGGCGATGCCATGAATTTCCTGGGCTACCG	1260
QY	1261	CCAGCTCAGATCTGAACAAGACAGAAACCTGTGTGGATCTTGTGTCTACTGTGAC	1320
Db	1261	CCAGCTCAGATCTGAACAAGACAGAAACCTGTGTGGATCTTGTGTCTACTGTGAC	1320
QY	1321	TGTCCCTGAGCAATTCACCTAAGAGGGTTGAGAAAGCTTTGCTGCCACCTGTGTCCAC	1380
Db	1321	TGTCCCTGAGCAATTCACCTAAGAGGGTTGAGAAAGCTTTGCTGCCACCTGTGTCCAC	1380
QY	1381	TCAGTCACTTTCCTGATGATGCTTCTGAGCCTTCCCTACATCTCTGAGCCTTAACATG	1440
Db	1381	TCAGTCACTTTCCTGATGATGCTTCTGAGCCTTCCCTACATCTCTGAGCCTTAACATG	1440
QY	1441	TCTGTGGGTATCACACAGATGTGAGTGTGTGTCACACAGTGTCTCAAGCAGAAAGACTTT	1500
Db	1441	TCTGTGGGTATCACACAGATGTGAGTGTGTGTCACACAGTGTGTCTCAAGCAGAAAGACTTT	1500
QY	1501	GTGTCCATGCTTGTCTAGAAAAACAGACTGGGGAACTTATGTAGACAGCACATCCAC	1560
Db	1501	GTGTCCATGCTTGTCTAGAAAAACAGACTGGGGAACTTATGTAGACAGCACATCCAC	1560
QY	1561	CAGTGAAGAACGGGATATGTCCTCTCTTCTGTATCTCTGTGATCTCTGTGCTGGGAGACTT	1620
Db	1561	CAGTGAAGAACGGGATATGTCCTCTCTTCTGTATCTCTGTGATCTCTGTGCTGGGAGACTT	1620
QY	1621	AGACTTTGTGGCCGTGAGAGGCTATTAAAGCAGCAGACAGATCAGTACAGTAATGATCATTA	1680
Db	1621	AGACTTTGTGGCCGTGAGAGGCTATTAAAGCAGCAGACAGATCAGTACAGTAATGATCATTA	1680
QY	1681	ACCTCCCTGTGCACATTTTGGCCCAATGGGGAAATGATCTTTACCAAGAGCTCACACGC	1740
Db	1681	ACCTCCCTGTGCACATTTTGGCCCAATGGGGAAATGATCTTTACCAAGAGAGCTCACACGC	1740
QY	1741	ATTTCACAGAGATGCAAAATCTGAGCCCTTGAGATTCGCCAGTGGATTCAAGAGGAA	1800
Db	1741	ATTTCACAGAGATGCAAAATCTGAGCCCTTGAGATTCGCCAGTGGATTCAAGAGGAA	1800
QY	1801	GTGGGAACAAAGGTGGATGCTCTATTAGAGCTTGACATCACAGCTATCGGTAATCAGA	1860
Db	1801	GTGGGAACAAAGGTGGATGCTCTATTAGAGCTTGACATCACAGCTATCGGTAATCAGA	1860
QY	1861	AATATGAAACAAATCTCTGACAAAAAGAGCAAGCTTAAATTTACAGAGGTGCTTGAGC	1920
Db	1861	AATATGAAACAAATCTCTGACAAAAAGAGCAAGCTTAAATTTACAGAGGTGCTTGAGC	1920
QY	1921	TGCATTGGAATATCACTTCCCTCTGCAATTTTCCATCAGATAGAAAGACTTGAACCTGTG	1980
Db	1921	TGCATTGGAATATCACTTCCCTCTGCAATTTTCCATCAGATAGAAAGACTTGAACCTGTG	1980
QY	1981	AAGCTGCCATCTGTTAATACTAAATTTCCCAATAGAAAAAATAGAAAAAATAGAAAAA	2032
Db	1981	AAGCTGCCATCTGTTAATACTAAATTTCCCAATAGAAAAAATAGAAAAAATAGAAAAA	2032
RESULT 2			
AAK94229			
ID AAK94229 standard; cDNA; 1979 BP.			
AAK94229;			
AC 06-NOV-2001 (first entry)			
XX Human full-length cDNA, SEQ ID NO: 2816.			
DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.			
XX			

OS Homo sapiens.
XX
XX EPI130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX P-PSDB: AAM93309.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX
XX Claim 8; SEQ ID NO 2816; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX
SQ Sequence 1979 BP; 451 A; 542 C; 474 G; 512 T; 0 other;

Query Match 88.7%; Score 1802; DB 22; Length 1979;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1902; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 114 ACTTTCACAGCTTCTGAGAGCAGAGTGTCTTCAAGCCGCTTTCAGAGCTTTCACCT 173
Db 1 ACTTTCACAGCTTCTGAGAGCAGAGTGTCTTCAAGCCGCTTTCAGAGCTTTCACCT 60

QY 174 TCAGACACATGCTACTGCTTAAAAAATGAAAGTCTGCTGTTTCTGTTTCCAGATGG 233
Db 61 TCAGACACATGCTACTGCTTAAAAAATGAAAGTCTGCTGTTTCTGTTTCCAGATGG 120

QY 234 CCATCTGGAGCTCTATCTTCTTCCACATGTAACAGCACAACATAGAGTCCCTGCTATGAAG 293
Db 121 CCACTCTGGAGCTCTATCTTCTTCCACATGTAACAGCACAACATAGAGTCCCTGCTATGAAG 180

QY 294 CACAGCCGAGCGAGATGACGTGCTGTTGTTGTTCTTCTGAGCGCTGAGCTTCTTTTG 353
Db 181 CACAGCCGAGCGAGATGACGTGCTGTTGTTGTTCTTCTGAGCGCTGAGCTTCTTTTG 240

QY 354 TGGGGGAGCTTTTGGGGAGCAGCCAGATGTTTCTTACCTATGAGACCCGCTGGACAG 413
Db 241 TGGGGGAGCTTTTGGGGAGCAGCCAGATGTTTCTTACCTATGAGACCCGCTGGACAG 300

QY 414 TGTGATGACCTTCAAGCAGAGCAGCCGCTGATGCTGACATAGGCTGTGGGATCTGA 473
Db 301 TGTGATGACCTTCAAGCAGAGCAGCCGCTGATGCTGACATAGGCTGTGGGATCTGA 360

QY 474 TACGGGCGCTTCTTCTGTTGGGACATGAGCGCTTGTGATGCTTAATGGAACCTGCTCCC 533
Db 361 TACGGGCGCTTCTTCTGTTGGGACATGAGCGCTTGTGATGCTTAATGGAACCTGCTCCC 420

QY 534 GGAGACAGTCCAGCCTTTTCACTGGGAGAACAGCCGGGCGCTGTGTTCTGACCTGCT 593

Db 421 GGAGACAGTCCAGCCTTTTCACTGGGAGAACAGCCGGGCGCTGTGTTCTGACCTGCT 480

QY 594 GTGACATCATCCCAAGAGATGAAATCATCCCCGGGGCTCACTGAGGCTCTTGTGACATC 653
Db 481 GTGACATCATCCCAAGAGATGAAATCATCCCCGGGGCTCACTGAGGCTCTTGTGACATC 540

QY 654 AACAGCCCTTTGAGAGTGTGGAGAGAGGCTGCGCTCTCTACAGCAGAGTGGTCTCAAG 713
Db 541 AACAGCCCTTTGAGAGTGTGGAGAGAGGCTGCGCTCTCTACAGCAGAGTGGTCTCAAG 600

QY 714 AGGTGGCTTCTTCACTGAGTCCCTCTTACCCGCTGCTGAAGACCCCTCCCTCAACC 773
Db 601 AGGTGGCTTCTTCACTGAGTCCCTCTTACCCGCTGCTGAAGACCCCTCCCTCAACC 660

QY 774 TGCATATCGTGCACCTGTCGGGGACCCCGGGCGGTGTCGTTCCCGGAGAACGACACAA 833
Db 661 TGCATATCGTGCACCTGTCGGGGACCCCGGGCGGTGTCGTTCCCGGAGAACGACACAA 720

QY 834 AGGAGATCTCATGATTGACAGTGCATTTGTATGGGGCAGATGAGCAGAAATCAGAGA 893
Db 721 AGGAGATCTCATGATTGACAGTGCATTTGTATGGGGCAGATGAGCAGAAATCAGAGA 780

QY 894 AGGAGACCAACCTTACTATGTATGTAGAGTCACTCTGCCAAAGCCAGCTGAGATCTACA 953
Db 781 AGGAGACCAACCTTACTATGTATGTAGAGTCACTCTGCCAAAGCCAGCTGAGATCTACA 840

QY 954 AGACATTCAGTCTTGGCCCAAGGCGCTGCAGAGAACGCTACCTGCTTGTGCTATGAG 1013
Db 841 AGACATTCAGTCTTGGCCCAAGGCGCTGCAGAGAACGCTACCTGCTTGTGCTATGAG 900

QY 1014 ACCGTGCTCAGGCGCTGTGAGCCCAAGTCCGAGATGTATGTAATTCGTGGATTTGGA 1073
Db 901 ACCGTGCTCAGGCGCTGTGAGCCCAAGTCCGAGATGTATGTAATTCGTGGATTTGGA 960

QY 1074 TCTTGCCCATCTTCAAGCTGGGTGATGATCAATCAACCCAGGCAAGGGATGGTGAGC 1133
Db 961 TCTTGCCCATCTTCAAGCTGGGTGATGATCAATCAACCCAGGCAAGGGATGGTGAGC 1020

QY 1134 ACCGTTTCCACACAATGCCAGGAGTCCCTTAATGTCTCCAGGCTTGGCGTGTCTT 1193
Db 1021 ACCGTTTCCACACAATGCCAGGAGTCCCTTAATGTCTCTCCAGGCTTGGCGTGTCTT 1080

QY 1194 TGGCCTATGAAAAGGTTTCTCGACTTGAAGAACCTGTGGCGATGCGATTAATTGCTGG 1253
Db 1081 TGGCCTATGAAAAGGTTTCTCGACTTGAAGAACCTGTGGCGATGCGATTAATTGCTGG 1140

QY 1254 GCTACGCGCAGTGCATGTGAAACAGAACAGAGAAACCTGTGTGATCTTGTGCTA 1313
Db 1141 GCTACGCGCAGTGCATGTGAAACAGAACAGAGAAACCTGTGTGATCTTGTGCTA 1200

QY 1314 CCTGAGCTGTCCCTGAGCAATCCACAAAGAGGTTGAGAAAGCTTTGCGCCACTGGT 1373
Db 1201 CCTGAGCTGTCCCTGAGCAATCCACAAAGAGGTTGAGAAAGCTTTGCGCCACTGGT 1260

QY 1374 GTAGGCTCAGTCACTTCTCTGAATGCTTCTAGGCTTCTCTACATCTCTGAGACCTTAA 1433
Db 1261 GTAGGCTCAGTCACTTCTCTGAATGCTTCTAGGCTTCTCTACATCTCTGAGACCTTAA 1320

QY 1434 CTACATGCTGTGGGTATACACATGAGTGTGATGTGTGCAACAGTGTCTCAAGCAGAA 1493
Db 1321 CTACATGCTGTGGGTATACACATGAGTGTGATGTGTGCAACAGTGTCTCAAGCAGAA 1380

QY 1494 GACTTTTGTGTCATGTTTGTCTAGAAAACAGACAGGGAACCTTAATGTAGAGACAG 1553
Db 1381 GACTTTTGTGTCATGTTTGTCTAGAAAACAGACAGGGAACCTTAATGTAGAGACAG 1440

QY 1554 ATCCACCAATGAAACAGGATATTGCTCTTCTTTCTTGAATCTTCTGCTGTGGGAG 1613
Db 1441 ATCCACCAATGAAACAGGATATTGCTCTTCTTTCTTGAATCTTCTGCTGTGGGAG 1500

QY 1614 ACTTCAGAGACTTGTGGCTGTGAGGCTTATTAAGACAGACAGATCATGATGGAATTGA 1673

DB 1501 ACTTCAGAGACTTGTGGCGCTGGAGGCTATTGAAGCAGACAGATATCATGGAATTGA 1560
QY 1674 TCCATTAACCTCCCTGTCACATCTTGGCCAAATGGGAAATGATCTTTACCAAGAAGCT 1733
DB 1561 TCCATTAACCTCCCTGTCACATCTTGGCCAAATGGGAAATGATCTTTACCAAGAAGCT 1620
QY 1734 CACCAGCATTTTCCACAGAGATGCAAAATTTCTGAGCCCTTGGAGTTCCAGTGATTTCAAG 1793
DB 1621 CACACACATTTTCCACAGAGATGCAAAATTTCTGAGCCCTTGGAGTTCCAGTGATTTCAAG 1680
QY 1794 GAAGGAAGTGGGAACAAGGTGGATGCTTACTTATGAGTTACACCTCAGCTATGCT 1853
DB 1681 GAAGGAAGTGGGAACAAGGTGGATGCTTACTTATGAGTTACACCTCAGCTATGCT 1740
QY 1854 AATCAGAAATATGAACAAATATCTGTCACAAAGAGAGAGCTTTTAACTTACAGAGGTG 1913
DB 1741 AATCAGAAATATGAACAAATATCTGTCACAAAGAGAGAGCTTTTAACTTACAGAGGTG 1800
QY 1914 CCGGGGCTGATTTGAATATCACTTCCCTGATTTCCATCATCATAGAAAGCTTTG 1973
DB 1801 CCGGGGCTGATTTGAATATCACTTCCCTGATTTCCATCATCATAGAAAGCTTTG 1860
QY 1974 ACCTGTGAAGCTGCATCTGTTAATCTAAATTTCCCAATTAAG 2017
DB 1861 ACCTGTGAAGCTGCATCTGTTAATCTAAATTTCCCAATTAAG 1904
RESULT 3
ID AA294211 standard; cDNA; 2065 BP.
AC AA294211;
XX
DT 19-JUN-2000 (first entry)
DE Human transferase TRNSFS-11 cDNA clone 2617407CBI.
XX
XX Transferase; TRNSFS-11; human; antitumor; cell proliferation;
KM inflammation; gastrointestinal disorder; developmental disorder;
KM genetic disorder; neurological disorder; reproductive disorder;
KM smooth muscle disorder; immunological disorder; gene therapy;
KM diagnosis; N-acetylglucosamine 6-O-sulfotransferase; ss.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 174..1334
FT /*tag= a
XX
XX MO200014251-A2.
XX
XX 16-MAR-2000.
XX
XX 09-SEP-1999; 99WC-US20989.
XX
XX 10-SEP-1998; 98US-0150657.
XX 04-NOV-1998; 98US-0186779.
XX 11-MAY-1999; 99US-0133642.
XX
XX (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
PI Hillman JL, Azimzai Y;
XX
XX WPI; 2000-256996/22.
XX
XX P-PSDB; AAY79219.
XX
XX Human transferase proteins useful for preventing, diagnosing and
XX treating cancers and developmental, gastrointestinal, genetic, and
XX immunological, neurological, reproductive and smooth muscle disorders -
XX Claim 9; Page 104-105; 113pp; English.

CC The present sequence is that of cDNA clone 2617407CBI encoding
CC human transferase TRNSFS-11 (see AAY79219), 1 of 15 claimed human
CC transferase proteins of the invention (see AAY79209-23). The clone
CC was isolated from gall bladder cDNA library GBLA0701. TRNSFS-11
CC is expressed in dermatologic and gastrointestinal tissues,
CC especially those associated with inflammation and cell
CC proliferation. It shows homology to mouse N-acetylglucosamine
CC 6-O-sulfotransferase. A polynucleotide comprising nucleotides
CC 264-333 or 1272-1331 of the present sequence can be used as a DNA
CC probe. The new human transferases and polynucleotides can be used
CC in the diagnosis, prevention and treatment (including gene therapy
CC and antisense therapy) of cancer, developmental disorders,
CC gastrointestinal disorders, genetic disorders, immunological
CC disorders, neurological disorders, reproductive disorders, and
CC smooth muscle disorders.
XX
XX
SQ Sequence 2065 BP; 466 A; 575 C; 491 G; 533 T; 0 other;
Query Match 79.5%; Score 1615; DB 21; Length 2065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2005; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 9 GCCAGGATGCTCCAGTCTGGGGGAAATGCTTCTATTTGCTTCCAGCCACCTC 68
DB 1 GCCAGGATGCTCCAGTCTGGGGGAAATGCTTCTATTTGCTTCCAGCCACCTC 60
QY 69 AAGCAGTCTCCCAACCCCTTGAAGTCTGACAGTGAAGCTTTTACAGCTTCC 128
DB 61 AAGCAGTCTCCCAACCCCTTGAAGTCTGACAGTGAAGCTTTTACAGCTTCC 120
QY 129 TGGAGCAGTGTCTTCTCAAGCCCGTCTTCAAGAGTCTTCCACTTCAAGCAATGCTAC 188
DB 121 TGGAGCAGTGTCTTCTCAAGCCCGTCTTCAAGAGTCTTCCACTTCAAGCAATGCTAC 180
QY 189 TGCCTAAATAATGAAGCTCTGCTGTTTCTGTTTCCCAATGAGCATCTGCTAT 248
DB 181 TGCCTAAATAATGAAGCTCTGCTGTTTCTGTTTCCCAATGAGCATCTGCTAT 240
QY 249 TCTTCCATGTATACAGCCCAACAATCAAGCTCCCTGTATGAAGCAGCCAGCGCA 308
DB 241 TCTTCCATGTATACAGCCCAACAATCAAGCTCCCTGTATGAAGCAGCCAGCGCA 300
QY 309 TGCAGTGTCTGTTCTGTTCTTCTGCGCTCTGCTCTTCTTTTGTGGGAGCTTTTG 368
DB 301 TGCAGTGTCTGTTCTGTTCTTCTGCGCTCTGCTCTTCTTTTGTGGGAGCTTTTG 360
QY 369 GGCAGCACCAGATTTTCTACAGATGAGAGCCCGCTGCGACGTGATGACTTCA 428
DB 361 GGCAGCACCAGATTTTCTACAGATGAGAGCCCGCTGCGACGTGATGACTTCA 420
QY 429 AGCAGAGCACCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 488
DB 421 AGCAGAGCACCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 489 TGTGGAATGAGTCTTGTATGCTGATGATGATGATGATGATGATGATGATGATGAT 548
DB 481 TGTGGAATGAGTCTTGTATGCTGATGATGATGATGATGATGATGATGATGATGAT 540
QY 549 TCTTCACTGGGAGAAACAGCCGGGCGCTGTTTGTGACCTGCTGTGATCATCTCCAC 608
DB 541 TCTTCACTGGGAGAAACAGCCGGGCGCTGTTTGTGACCTGCTGTGATCATCTCCAC 600
QY 609 AAGATGAATATATCCCGGGCTCAGTCAAGCTCTGTGATGATCAACAGCCCTTTG 667
DB 601 AAGATG-AATCATCCCGGGCTCAGTCAAGCTCTGTGATGATCAACAGCCCTTTG 659
QY 668 GTGTGAGAGAGCCCTGCGCTCTACAGCAGTGTGATGATGATGATGATGATGATGAT 727
DB 660 GTGTGAGAGAGCCCTGCGCTCTACAGCAGTGTGATGATGATGATGATGATGATGAT 719
QY 728 AACCTGAGTCTCTTACCCGCTGCTGAAGAACCCCTTCTCAACTGATATGATGATGAT 787
DB 720 AACCTGAGTCTCTTACCCGCTGCTGAAGAACCCCTTCTCAACTGATATGATGATGAT 779

QY	788	CTGTTCGGGAGACCCCGCGCCGCTTCCCTCCCGAGAAAGCACAAGAGGAAATCTCATG	847
Db	780	CTGTTCGGGAGACCCCGCGCGCGTTCCTCCGTCGAGAAAGCACAAGAGGAAATCTCATG	839
QY	848	ATTGACACTCGCATTTGTATGGGACAGCATGAGCAGAAACTCAAGAAGAGGACCAACCC	907
Db	840	ATTGACACTCGCATTTGTATGGGAGGAGCATGAGCAGAAACTCAAGAAGAGGACCAACCC	899
QY	908	TACTATGTGATGCGAGGTCATCTGCGAAAGCGACCTGGAATCTACAAAGCATCCAGTC	967
Db	900	TACTATGTGATGCGAGGTCATCTGCGAAAGCGACCTGGAATCTACAAAGCATCCAGTC	959
QY	968	TTTGCCCAAGGCCCTGACAGGAAGCTACCTGCTGTGCGATGAGAGCACTGGCTGAGCC	1027
Db	960	TTTGCCCAAGGCCCTGACAGGAAGCGTACCTGCTGTGCGATGAGAGCACTGGCTGAGCC	1019
QY	1028	CCTGTGGGCGCAGACTTCCCGCAATGTATGAAATTCGTGGGATTTGGAAATCTTGGCCATCT	1087
Db	1020	CCTGTGGGCGCAGACTTCCCGCAATGTATGAAATTCGTGGGATTTGGAAATCTTGGCCATCT	1079
QY	1088	CAGACTGGGTGCTAATACATCACCCGAGGCAAGGGCATGGGTGACACAGCTTCCACACA	1147
Db	1080	CAGACTGGGTGCTAATACATCACCCGAGGCAAGGGCATGGGTGACACAGCTTCCACACA	1139
QY	1148	AATGCGAGGGGATGGCCCTTATGTCTCCAGGCGCTGGCGCTGGCTTGGCCATATGAAG	1207
Db	1140	AATGCGAGGGATGGCCCTTATGTCTCCAGGCGCTGGCGCTGGCTTGGCCATATGAAG	1199
QY	1208	GTTTCTGCACTTCAGAAAGCCTGTGGCGATGCCATGAATTTCTGGGCTACCCGCAATC	1267
Db	1200	GTTTCTGCACTTCAGAAAGCCTGTGGCGATGCCATGAATTTCTGGGCTACCCGCAATC	1259
QY	1268	AGATCTCAACAGAAAGAGAAACCTGTGTGGATCTTCTGTACCTGGAGATCTGCCCT	1327
Db	1260	AGATCTCAACAGAAAGAGAAACCTGTGTGGATCTTCTGTACCTGGAGATCTGCCCT	1319
QY	1328	GAGCAAAATCCACTAAGAGGGTTAGAAAGGCTTGCTGCACCTGGTGTGACGCTTAGTCA	1387
Db	1320	GAGCAAAATCCACTAAGAGGGTTAGAAAGGCTTGCTGCACCTGGTGTGACGCTTAGTCA	1379
QY	1388	CTTCTCTCGAATGGTTGTGAGCCCTTGCATACATCTGTGAGCCTTAACATAGTCTGTGG	1447
Db	1380	CTTCTCTCGAATGGTTGTGAGCCCTTGCATACATCTGTGAGCCTTAACATAGTCTGTGG	1439
QY	1448	GTAATCACTGATGTGAGTGTGATGTGTCTCCACAGTGTCTAAGCAGAAAGCACTTTGTGTCA	1507
Db	1440	GTAATCACTGATGTGAGTGTGATGTGTCTCCACAGTGTCTAAGCAGAAAGCACTTTGTGTCA	1499
QY	1508	TGCTTTGTGTAGAAAACACACAGTGGGGAACCTTATGTGAGCGACACATCCACAGATGA	1567
Db	1500	TGCTTTGTGTAGAAAACACACAGTGGGGAACCTTATGTGAGCGACACATCCACAGATGA	1559
QY	1568	ACAGGGATTTGCTCTCTCTTTCTTTGATCTCTCTGCTGTGGGCGACTTCAGAGACTTT	1627
Db	1560	ACAGGGATTTGCTCTCTCTCTTTCTTTGATCTCTCTGCTGTGGGCGAGACTTCAGAGACTTT	1619
QY	1628	GTGGCTCTGGAGGCCATTATAGCACGACACAGATAGTGGAAATTAATCAATTAACCTGCC	1687
Db	1620	GTGGCTCTGGAGGCCATTATAGCACGACACAGATAGTGGAAATTAATCAATTAACCTGCC	1679
QY	1688	TGTCACACATTTGGCCCAATGGGGAATGATATCTTTACCAAGAGCTCACAGCAATTTTCC	1747
Db	1680	TGTCACACATTTGGCCCAATGGGGAATGATATCTTTACCAAGAGCTCACAGCAATTTTCC	1739
QY	1748	ACAGAGATGCAAAATTCGAGACCCCTTGAGATCTCCAGTGGATTCAAGAGAAGATGGGAA	1807
Db	1740	ACAGAGATGCAAAATTCGAGACCCCTTGAGATCTCCAGTGGATTCAAGAGAAGATGGGAA	1799
QY	1808	CAAGGTTGATGCTCTACTTATGAGCTTGACATCACAGCTATCGGTAATCAAGAAATATGA	1867
Db	1800	CAAGGTTGATGCTCTACTTATGAGCTTGACATCACAGCTATCGGTAATCAAGAAATATGA	1859

OY	1868	AAAAAATCTGTGACAAAGAGCAAGCCTTGTAGTCACAGAGTGCCCTGGCATTT	1927
XX			
DB	1860	AACAATATCTCTGCACAAAGAGCAAGCCTTTAAGTTCAAGAGTGCCCGGCTCATTT	1919
XX			
OY	1928	GAAATACACTTCCCCTGTGCATTTCCTCCATACATAGAGACTTTGACCTGTGAAGCTGC	1987
DB	1920	GAATATCACTTCCCCCTGTGCATTTCCTCCATACATAGAGACTTTGACCTGTGAAGCTGC	1979
OY	1988	CATCGTTTAATACATAAATTCGCCAATAAG	2017
DB	1980	CATCGTTTAATACATAAATTCGCCAATAAG	2009
XX			
XX			
ID	ASL16947	standard; cDNA; 1333 BP.	
AC	ASL16947;		
XX			
DT	12-MAR-2002	(first entry)	
XX			
DE	Human L-selectin sulfotransferase-2 (LSST-2) cDNA.		
XX			
KW	Human; beta1,3gnt; beta1,3-N-acetylglucosaminyltransferase; MECA-79; ss;		
KW	L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;		
KW	ulcerative colitis; inflammatory skin disorder; psoriasis; lichen planus;		
KW	allergic contact dermatitis; lymphoma; chronic pneumonia; LSST-2;		
KW	delayed-type hypersensitivity reaction; hyperplastic thymus; antidiacer;		
KW	antiinflammatory; antipsoriatic; antidiabetic; dermatological;		
KM	antiallergic.		
XX			
OS	Homo sapiens.		
XX			
FH	Location/Qualifiers		
FT	111..1253		
CD	/tag= a		
DS	/product= "Human LSST-2"		
ET			
FT			
XX			
FN	WO200185177-A1.		
PD	15-NOV-2001.		
XX			
PE	10-MAY-2001; 2001MO-US15452.		
XX			
PR	11-MAY-2000; 2000US-0569320.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Fukuda M, Yeh J, Hiraoka N;		
DR	WPI; 2002-075226/10.		
P-	P-PSDB; AA011274.		
PT	New enzyme, useful for modifying acceptor molecule, comprises an		
PT	isolated L-selectin sulfotransferase-2 that directs expression of		
PT	L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or		
PT	intestinal GlcNAc 6-sulfotransferase		
XX			
PS	ClaIm 19; Fig 4; 98pp; English.		
XX			
CC	The present invention provides a method of modifying an acceptor molecule		
CC	by contacting the acceptor with an isolated		
CC	beta1,3-N-acetylglucosaminyltransferase (beta1,3gnt) or an active		
CC	fragment, where beta1,3gnt directs expression of a MECA-79 antigen. The		
CC	invention also provides a method of treating or preventing an		
CC	L-selectin-mediated condition by reducing the expression or activity of a		
CC	beta1,3gnt that directs expression of a MECA-79 antigen. This can be done		
CC	by administering to the subject an oligosaccharide L-selectin antagonist		
CC	that inhibits binding of L-selectin to a MECA-79 antigen, for example by		
CC	administering antibody material that specifically binds beta1,3gnt,		
CC	and/or a beta1,3gnt antisense nucleic acid molecule. L-selectin		
CC	sulfotransferase-2 (LSST-2) also directs MECA-79 antigen expression.		
CC	Alternatively, the expression or activity of LSST-2 or its active		

CC fragmentation can be reduced in combination with reducing the expression or activity of heparanase. The method is useful for treating L-selectin mediated conditions such as Crohn's disease and ulcerative colitis, CC inflammatory disorders of the skin such as allergic contact dermatitis, CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type CC hypersensitivity reactions, diabetes and hyperplastic thymus. This CC sequence represents cDNA encoding human LST-2.

Sequence 1333 BP; 286 A; 393 C; 336 G; 318 T; 0 other;

Query Match 59.6%; Score 1211; DB 24; Length 1333;

```
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	161	AGGTCCTTCACCTTCAGCAAAATGCTACTGCTTAAAAAATTAACCTCTCTGTTCGTG	220
Db	72	AAGTCCTTCACCTTCAGCAAAATGCTACTGCTTAAAAAATTAACCTCTCTGTTCGTG	131
QY	221	GTTTCCAGATGGCATCTTGGCTCTATTCTTCACATGATACAGCCACAACATCACTGCC	280
Db	132	GTTTCCAGATGGCATCTTGGCTCTATTCTTCACATGATACAGCCACAACATCACTGCC	191
QY	281	CTGTCTATGAAGCAGCAAGCCCGAGCCATGACAGTGGCTGTCTCTCCCTGGCGCTCT	340
Db	192	CTGTCTATGAAGCAGCAAGCCCGAGCCATGACAGTGGCTGTCTCTCCCTGGCGCTCT	251
QY	341	GGCCTCTCTTTTGTGGGCGACGTTTTTGGGCGACACCCAGATGTTTTCTACCTGATGGAG	400
Db	252	GGCCTCTCTTTTGTGGGCGACGTTTTTGGGCGACACCCAGATGTTTTCTACCTGATGGAG	311
QY	401	CCCGCTGGCACCTGTGGATGACCTTCAAGCAGAGACCGCCTGGATCCTGCACATGGCT	460
Db	312	CCCGCTGGCACCGTGGATGACCTTCAAGCAGAGACCGCCTGGATCCTGCACATGGCT	371
QY	461	GTTGGGGATCTGATACGGGCGCTTCTCTGTGCGAGACATGAGGCTTTGATGGCCATAGT	520
Db	372	GTTGGGGATCTGATACGGGCGCTTCTCTGTGCGAGACATGAGGCTTTGATGGCCATAGT	431
QY	521	GAACCTGTGTCCTGGAGACAGTCCACGCTTCTTCATGTGGAGAACAGCCGGGCGCTGTGT	580
Db	432	GAACCTGTGTCCTGGAGACAGTCCACGCTTCTTCATGTGGAGAACAGCCGGGCGCTGTGT	491
QY	581	TCTGCACCTGCTGTGACATCATCCGACAAAGMTGAATATGCCCGGGGCTACATCGCAGG	640
Db	492	TCTGCACCTGCTGTGACATCATCCGACAAAGMTGAATATGCCCGGGGCTACATCGCAGG	551
QY	641	CTCCTGTGACAGTCAACAGCCCTTTGAAGGTGTGGAGAAAGGCTGCGGCTCTTACAGCCAC	700
Db	552	CTCCTGTGACAGTCAACAGCCCTTTGAAGGTGTGGAGAAAGGCTGCGGCTCTTACAGCCAC	611
QY	701	GTTGGTGTCAAGGAGGTGGGCTCTTTCACATCCGACAGTCCCTTACCGGTGTGTAAGAC	760
Db	612	GTTGGTGTCAAGGAGGTGGGCTCTTTCACATCCGACAGTCCCTTACCGGTGTGTAAGAC	671
QY	761	CCCTCCCTCAACCTGCATATGTGACACTGTGTCGGGAGCCCCGGGAGCCGTGTTCCTGTC	820
Db	672	CCCTCCCTCAACCTGCATATGTGACACTGTGTCGGGAGCCCCGGGAGCCGTGTTCCTGTC	731
QY	821	CGAACAAGCAAAAGGAGATCTCATGATTGACAGTGCAGTGTGATGGGAGCATAG	880
Db	732	CGAACAAGCAAAAGGAGATCTCATGATTGACAGTGCAGTGTGATGGGAGCATAG	791
QY	881	CAGAACTCAAGAGAGAGACCAACCTTACTATGTGATGCAGGTCATCTGCCAAGCCAG	940
Db	792	CAGAACTCAAGAGAGAGACCAACCTTACTATGTGATGCAGGTCATCTGCCAAGCCAG	851
QY	941	CTGAGAGTTCACAGACATCCAGTCTCTTGCCCAAGGCCCTGCAGAAAGCTACTCTGTT	1000
Db	852	CTGAGAGTTCACAGACATCCAGTCTCTTGCCCAAGGCCCTGCAGAAAGCTACTCTGTT	911
QY	1001	GTCGCGTATGAGGACTGCTGCTGCAGCCCTGTGTCGGCAGACTTCCGAATGTATGAATC	1066
Db	912	GTCGCGTATGAGGACTGCTGCTGCAGCCCTGTGTCGGCAGACTTCCGAATGTATGAATTC	971

QY	1061	GTGGATTTGGAATCTTGGCCCATCTTCCAGACTGGGTGCATTAATACCCGAGGCAAG	1120
Db	972	GTGGGATTTGGAATCTTGGCCCATCTTCCAGACTGGGTGCATTAATACCCGAGGCAAG	1031
QY	1121	GGCATGGGTGACACACGCTTTCACACAATAATGGCAGGAGATGCCCTTAATGTCTCCAGGCT	1180
Db	1032	GGCATGGGTGACACACGCTTTCACACAATAATGGCAGGAGATGCCCTTAATGTCTCCAGGCT	1091
QY	1181	TGSCGCTGGTCTTGTGCCCTATGAAAAGGTTTCTCGACTTCAGAAAGCCTGTGGGAGTCC	1240
Db	1092	TGGCGCTGGTCTTGTGCCCTATGAAAAGGTTTCTCGACTTCAGAAAGCCTGTGGGAGTCC	1151
QY	1241	ATGAATTTGCGGGGCTACGGCCACAGTCAGATCGAACAAGACAGACAGAAACCTGTGCTG	1300
Db	1152	ATGAATTTGCTGGGCTACGGCCACAGTCAGATCGAACAAGACAGACAGAAACCTGTGCTG	1211
QY	1301	GATCTTCTGCTACTGAGACTGTCCCTGAGCAAAATCCACTAAGAGGGTTGAGAAAGCCTT	1350
Db	1212	GATCTTCTGCTACTGAGACTGTCCCTGAGCAAAATCCACTAAGAGGGTTGAGAAAGCCTT	1271
QY	1361	GCTGCCACCTGCTGCTAGGCTCAGTCACTTCTCTAATAGCTTTCAGAGCCTTGCCCTACAT	1420
Db	1272	GCTGCCACCTGCTGCTAGGCTCAGTCACTTCTCTAATAGCTTTCAGAGCCTTGCCCTACAT	1331
QY	1421	CT 1422	
Db	1332	CT 1333	

RESULT 5
AAK91803

AC AAK918037

DT 06-NOV-2001 (first entry)

DE Human cDNA 5'-end sequence, SEQ ID NO: 263.

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EP1130094-A2

PD 05-SEP-2001

PF 07-JUL-2000; 2000EP-0114089

PR 08-JUL-1999; 99JP-0194486

PR 02-MAY-2000; 2000JP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX
DD
FMT: 2001-EOAC/EE/EO

XX
XX
gag
for
cDNA
and
their

PS Claim 2; SEQ ID NO 263; 1380bp + sequence listing; English

CC The invention relates to primers for synthesising full length cDNA
CC clones. 380 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 5'-end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 877 BP; 183 A; 266 C; 221 G; 203 T; 4 other;

Query Match 36.5%; Score 741; DB 22; Length 877;
 Best Local Similarity 100.0%; Pred. No. 1,1e-305; Mismatches 0; Indels 0; Gaps 0;
 Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 ACTTTCACAGCTTCTGAGAGCGAGTCTTCTCAAGCCGCTCTTGCAGAGTCTTCCACT 173
 DB 1 ACTTTCACAGCTTCTGAGAGCGAGTCTTCTCAAGCCGCTCTTGCAGAGTCTTCCACT 60
 QY 174 TCAGCAGATGCTAGTCCCTAAAAAATGAAGCTCTGCTGCTTGTGTTTCCAGATGG 233
 DB 61 TCAGCAGATGCTAGTCCCTAAAAAATGAAGCTCTGCTGCTTGTGTTTCCAGATGG 120
 QY 234 CCATCTTGCTCTATCTTCCACATGTACAGCCACACATCAGTCCGCTCTATGAAG 293
 DB 121 CCATCTTGCTCTATCTTCCACATGTACAGCCACACATCAGTCCGCTCTATGAAG 180
 QY 294 CACAGCCGAGGCGATGACGCTGCTGCTCTCTTCTGCTGCTGCTGCTGCTGCTT 353
 DB 181 CACAGCCGAGGCGATGACGCTGCTGCTCTCTTCTGCTGCTGCTGCTGCTGCTT 240
 QY 354 TGGGCGAGCTTTTGGGAGAGCACCAGATGTTTCTACCTGATGAGCCGCTGGACG 413
 DB 241 TGGGCGAGCTTTTGGGAGAGCACCAGATGTTTCTACCTGATGAGCCGCTGGACG 300
 QY 414 TGTGATGACCTTCAAGACAGACACCCGCTGATCTGACATGGCTCTGCGGATCGA 473
 DB 301 TGTGATGACCTTCAAGACAGACACCCGCTGATCTGACATGGCTCTGCGGATCGA 360
 QY 474 TACGGCCGCTCTTCTGTCGACATGAGCGTCTTGTGATGCTGATGAGACCTGATCCC 533
 DB 361 TACGGCCGCTCTTCTGTCGACATGAGCGTCTTGTGATGCTGATGAGACCTGATCCC 420
 QY 534 GGAGACAGTCCAGCCTCTTTCAGTGGGAGAACAGCCGCGCTGCTGCTGACCTGCT 593
 DB 421 GGAGACAGTCCAGCCTCTTTCAGTGGGAGAACAGCCGCGCTGCTGCTGACCTGCT 480
 QY 594 GTGACATATCCCAAGATGAATCATCCCGGCTCAGTGGAGCTCTGCTGCTGCTGCT 653
 DB 481 GTGACATATCCCAAGATGAATCATCCCGGCTCAGTGGAGCTCTGCTGCTGCTGCT 540
 QY 654 AACAGCCCTTGTAGGTGTGAGAAAGGCTGCTGCTCTACAGCAGCAGTGTGCTCAAG 713
 DB 541 AACAGCCCTTGTAGGTGTGAGAAAGGCTGCTGCTCTACAGCAGCAGTGTGCTCAAG 600
 QY 714 AGTGGCGCTTCTTCAACTGCACTGCACTCTTACCCGCTCTGAAAGACCCCTTCCCAAC 773
 DB 601 AGTGGCGCTTCTTCAACTGCACTGCACTCTTACCCGCTCTGAAAGACCCCTTCCCAAC 660
 QY 774 TGCATATGCTGACCTGCTGCTGCGGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCT 833
 DB 661 TGCATATGCTGACCTGCTGCTGCGGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 834 AGGAGATCTCATGATTACA 854
 DB 721 AGGAGATCTCATGATTACA 741

RESULT 6
 AAK93921
 ID AAK93921 standard; cDNA: 877 BP.

XX AAK93921;
 AC AAK93921;
 XX 06-NOV-2001 (first entry)
 DT
 XX

DE Human cDNA clone representative sequence, SEQ ID NO: 2381.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 KM Homo sapiens.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Homo sapiens.
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 XX 07-JUL-2000; 2000EP-0114089.
 PF
 XX 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 PI WPI: 2001-52425/58.
 PT
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 PS
 XX Example 11; SEQ ID NO 2381; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' and 3' ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence was used as the
 CC representative sequence from a human clone which was used in
 CC homology searches to identify the clone.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

SO Sequence 877 BP; 183 A; 266 C; 221 G; 203 T; 4 other;

Query Match 36.5%; Score 741; DB 22; Length 877;
 Best Local Similarity 100.0%; Pred. No. 1,1e-305; Mismatches 0; Indels 0; Gaps 0;
 Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 ACTTTCACAGCTTCTGAGAGCGAGTCTTCTCAAGCCGCTCTTGCAGAGTCTTCCACT 173
 DB 1 ACTTTCACAGCTTCTGAGAGCGAGTCTTCTCAAGCCGCTCTTGCAGAGTCTTCCACT 60
 QY 174 TCAGCAGATGCTAGTCCCTAAAAAATGAAGCTCTGCTGCTTGTGTTTCCAGATGG 233
 DB 61 TCAGCAGATGCTAGTCCCTAAAAAATGAAGCTCTGCTGCTTGTGTTTCCAGATGG 120
 QY 234 CCATCTTGCTCTATCTTCCACATGTACAGCCACACATCAGTCCGCTCTATGAAG 293
 DB 121 CCATCTTGCTCTATCTTCCACATGTACAGCCACACATCAGTCCGCTCTATGAAG 180
 QY 294 CACAGCCGAGGCGATGACGCTGCTGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCT 353
 DB 181 CACAGCCGAGGCGATGACGCTGCTGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 354 TGGGCGAGCTTTTGGGAGAGCACCAGATGTTTCTACCTGATGAGCCGCTGGACG 413
 DB 241 TGGGCGAGCTTTTGGGAGAGCACCAGATGTTTCTACCTGATGAGCCGCTGGACG 300
 QY 414 TGTGATGACCTTCAAGACAGACACCCGCTGATCTGACATGGCTCTGCGGATCGA 473
 DB 301 TGTGATGACCTTCAAGACAGACACCCGCTGATCTGACATGGCTCTGCGGATCGA 360
 QY 474 TACGGCCGCTCTTCTGTCGACATGAGCGTCTTGTGATGCTGATGAGACCTGATCCC 533

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 3; SEQ ID NO 1048; 1380bp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 nucleotide sequences encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 3'-end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 XX
 XX Sequence 548 BP; 160 A; 109 C; 129 G; 148 T; 2 other;

Query Match 17.3%; Score 351; DB 22; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1.5e-139;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1667 GATTGATCCATTAACCTCCCTGTCACATCTTGGCCATGGGGAATGGATCTTACCA 1726
 DB 426 GATTGATCCATTAACCTCCCTGTCACATCTTGGCCATGGGGAATGGATCTTACCA 367
 OY 1727 TAGACTACCCGATTTTCCACAGATGCAATCTGAGCCCTTGGAGTTCCCAATGG 1786
 DB 366 TAGACTACCCGATTTTCCACAGATGCAATCTGAGCCCTTGGAGTTCCCAATGG 307
 OY 1787 ATTCAAG 1846
 DB 306 ATTCAAG 247
 OY 1847 TTTCGGTATGCAATTAATGAACAAATCTGTGACAAAGAGAGAGAGAGAGAGAG 1906
 DB 246 TTTCGGTATGCAATTAATGAACAAATCTGTGACAAAGAGAGAGAGAGAGAGAG 187
 OY 1907 GAGGCTGCTGGGCTGATTTGAATATCACTTCCCTGATTTCCCATCATAGAA 1966
 DB 186 GAGGCTGCTGGGCTGATTTGAATATCACTTCCCTGATTTCCCATCATAGAA 127
 OY 1967 GACTTTGACCTTGAACTGCGCATCTGTTAATCTAAATTCCTCAATTAAG 2017
 DB 126 GACTTTGACCTTGAACTGCGCATCTGTTAATCTAAATTCCTCAATTAAG 76

RESULT 9
 AAC76156
 ID AAC76156 standard; cDNA; 2988 BP.

AC AAC76156;

DE 08-FEB-2001 (first entry)

XX Human ORF1711 polynucleotide sequence SEQ ID NO:3421.

XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
 KW vulnery; antiparasitic; antiparkinsonian; neuroprotective;
 KW anticonvulsant; osteoporotic; antiparkinsonian; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antiinflammatory;
 KW hypotensive; dermatologic; immunosuppressive; antineoplastic; antitumor;
 KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200058473-A2.
 XX
 XX 05-Oct-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 XX 02-APR-1999; 99US-0127636.
 XX 05-APR-1999; 99US-0127728.
 XX 30-MAR-2000; 2000US-0540763.
 XX
 XX (CUTRA-) CURAGEN CORP.
 XX
 XX Shinketsu RA, Leach M;
 PI
 XX WPI; 2000-602362/57.
 XX P-PSDB; AAB41947.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 2597-2599; 5507pp; English.

AACT4446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,
 which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; neuroprotective;
 CC osteoporotic; anticonvulsant; antiparkinsonian; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatologic; immunosuppressive;
 CC antineoplastic; antibacterial; antiviral; antineoplastic; antithyroid;
 CC antianemic; antiparasitic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SO Sequence 2988 BP; 613 A; 916 C; 842 G; 617 T; 0 other;

Query Match 15.1%; Score 306; DB 21; Length 2988;
 Best Local Similarity 100.0%; Pred. No. 2e-120;

Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 506 TTGATGCTATGAGAACTGCTGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 565
 DB 1480 TTGATGCTATGAGAACTGCTGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
 OY 566 AGCCGGGCGCTGTGTGTGACCTGCTGACATCTCCCAAGATGAATGATCC 625
 DB 1540 AGCCGGGCGCTGTGTGTGACCTGCTGACATCTCCCAAGATGAATGATCC 1599
 OY 626 CGGCTACAGCAGAGCTCTGTGACATCAACAGCCCTTTGAGGTGTGAGAGAGCTGC 685
 DB 1600 CGGCTACAGCAGAGCTCTGTGACATCAACAGCCCTTTGAGGTGTGAGAGAGCTGC 1659

CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 60 BP; 13 A; 20 C; 9 G; 18 T; 0 other;
 SQ

Query Match 3.0%; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1934 CACTTCCCTGTCATTTCCATCAGAGAACTTGACCTGTGAGCTGCATCTG 1993
 DB 1 CACTTCCCTGTCATTTCCATCAGAGAACTTGACCTGTGAGCTGCATCTG 60

RESULT 12
 AAD24670
 ID AAD24670 standard; CDNA; 1647 BP.
 AC AAD24670;
 XX
 XX 12-MAR-2002 (first entry)
 DE Human drug metabolizing enzyme (DME)-5 CDNA.
 XX
 XX Human; drug metabolizing enzyme; gene therapy; autoimmune disorder;
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;
 KW cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;
 KW DME-5; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1.1188
 FT /tag= a
 FT /product= "Human DME-5 protein"
 FT sig_peptide 1..96
 FT /tag= b
 FT mat_peptide 97..1185
 FT /tag= c
 FT /product= "Human mature DME-5 protein #1"
 FT sig_peptide 1..105
 FT /tag= d
 FT mat_peptide 106..1185
 FT /tag= e
 FT /product= "Human mature DME-5 protein #2"
 XX
 XX WO2001/9468-A2.
 PN
 XX 25-OCT-2001.
 PD
 XX 12-APR-2001; 2001WO-US11869.
 PF
 XX 13-APR-2000; 2000US-197590P.
 PR 19-APR-2000; 2000US-196403P.
 PR 28-APR-2000; 2000US-200185P.
 PR 05-MAY-2000; 2000US-202234P.
 PR 11-MAY-2000; 2000US-203509P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Policky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM,
 PI Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS,
 PI Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 PI Au-Young J;
 XX

DR WPI: 2002-066363/09.
 DR P-PSDB: AAB15438.
 XX
 XX Novel isolated human drug metabolizing enzymes referred as DME 1-10,
 FT useful for diagnosing, treating, or preventing disorders associated
 PT with aberrant expression of DME such as allergy, anemia, asthma,
 XX infertility
 XX
 XX Claim 5; Page 139; 143pp; English.
 PS
 XX The invention relates to human drug metabolizing enzymes referred as
 CC DME and nucleic acid molecules encoding such enzymes. Polynucleotides
 CC of the invention are useful for assessing toxicity of test compounds
 CC and in gene therapy. Sequences of the invention are useful in the
 CC diagnosis, prevention and treatment of autoimmune/inflammatory
 CC disorders such as acquired immune deficiency syndrome (AIDS), adult
 CC respiratory distress syndrome, allergies, anaemia, arteriosclerosis,
 CC asthma, autoimmune haemolytic anaemia, contact dermatitis, Crohn's
 CC disease, glomerulonephritis, Goodpasture's syndrome, gout, Grave's
 CC disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple
 CC sclerosis, osteoarthritis, osteoporosis, psoriasis, systemic lupus
 CC erythematosus, rheumatoid arthritis, scleroderma, ulcerative colitis,
 CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections; cell proliferative disorders such as actinic keratosis,
 CC arteriosclerosis, atherosclerosis, Duchenne and Becker dystrophy,
 CC hepatitis, Cushing's syndrome, cancers, myelodysplastic syndrome,
 CC epilepsy; endocrine disorders such as disorders of the hypothalamus
 CC and pituitary resulting from lesions such as primary brain tumours,
 CC adenomas, infarction associated with pregnancy, aneurysms, vascular
 CC malformations; eye disorders such as conjunctivitis, iritis, retinitis,
 CC glaucoma, pigmentosa; metabolic disorders such as Addison's disease,
 CC cystic fibrosis, diabetes, goitre, glycogen storage diseases,
 CC hypercholesterolaemia, hyperthyroidism, hypoglycaemia, lipid myopathies,
 CC Meckles syndrome, mannosidosis, obesity; gastrointestinal disorders such
 CC as dysphagia, gastric carcinoma, anorexia, nausea, gastroenteritis,
 CC hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's
 CC syndrome, peliosis hepatitis, hepatic vein thrombosis and developmental
 CC disorders. The present sequence is human DME-5 protein CDNA.
 CC
 XX
 XX Sequence 1647 BP; 287 A; 560 C; 501 G; 299 T; 0 other;
 SQ

Query Match 2.9%; Score 59; DB 24; Length 1647;
 Best Local Similarity 100.0%; Pred. NO. 3.4e-15;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGGCTCCTACAGCCAGCGTGTGTCAGAGAGTGGCTTTCACCTGCAG 736
 DB 488 AGGCTGCGGCTCCTACAGCCAGCGTGTGTCAGAGAGTGGCTTTCACCTGCAG 546

RESULT 13
 AAD027700
 ID AAD027700 standard; CDNA; 1694 BP.
 AC AAD027700;
 XX
 XX 02-MAY-2001 (first entry)
 DE Human glycosyl sulfotransferase-4beta (GST-4beta) CDNA.
 XX
 XX Human; glycosyl sulfotransferase-4beta; GST-4beta; immunosuppressive;
 KW therapy; selectin binding inhibitor; gene therapy; inflammation;
 KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 KW glomerulonephritis; mastenena gravis; Sjogren's syndrome; adrenailitis;
 KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
 KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
 KW chromosome 16q23.1; ss.
 XX
 XX Homo sapiens.
 OS
 XX

```

FH Key Location/Qualifiers
FT 5'UTR 9..188
FT /*tag= a
FT CDS 189..1376
FT /product= "Human glycosyl sulfotransferase-4beta
FT /note= "CDS is referred as SEQ ID NO:21 in brief
FT description of the figures (page no: 4)"
FT 3'UTR 1377..1694
FT /*tag= c
XX
XX MO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX (RBGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX WPI: 2001-138471/14.
XX P-PSDB; AAY72640.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Claim 6; Fig 4A; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4beta (GST-4
XX beta) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation.
XX
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX Sequence 1694 BP; 259 A; 586 C; 514 G; 335 T; 0 other;
XX
XX Query Match 2.9%; Score 59; DB 22; Length 1694;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-15;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 678 AGGCGTGGCGCTCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCGCTTCTCAACCTGCAG 736
XX |||||||
XX Db 676 AGGCGTGGCGCTCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCGCTTCTCAACCTGCAG 734
XX |||||||

```

```

XX therapy; selectin binding inhibitor; gene therapy; inflammation;
XX systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
XX polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
XX glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
XX Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
XX demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
XX myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
XX asthma; hypersensitivity; rheumatic fever; tissue rejection;
XX chromosome 16q23.1; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 218..1390
XX /*tag= a
XX /product= "Human glycosyl sulfotransferase-4alpha
XX (GST-4alpha)"
XX /note= "CDS is specifically claimed as SEQ ID NO: 4
XX in claim 6 (page no: 41) of the specification"
XX
XX MO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX (RBGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX WPI: 2001-138471/14.
XX P-PSDB; AAY72639.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Claim 6; Fig 1; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
XX alpha) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation.
XX
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX Note: The present sequence is also shown in sequence listing (page
XX no: 56) but lacks four nucleotides at its 3' end.
XX
XX Sequence 2044 BP; 422 A; 645 C; 570 G; 407 T; 0 other;
XX
XX Query Match 2.9%; Score 59; DB 22; Length 2044;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-15;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 678 AGGCGTGGCGCTCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCGCTTCTCAACCTGCAG 736
XX |||||||
XX Db 708 AGGCGTGGCGCTCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCGCTTCTCAACCTGCAG 766
XX |||||||

```

RESULT 15

AAD02698
ID AAD02698 standard; cDNA: 2170 BP.
XX
AC AAD02698;
XX
DT 02-MAY-2001 (first entry)
DE Human glycosyl sulfotransferase-4alpha (GST-4alpha) cDNA #1.
XX
KW Human: glycosyl sulfotransferase-4alpha; GST-4alpha; immunosuppressive;
KW therapy: selectin binding inhibitor; gene therapy; inflammation;
KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
KW chromosome 16q23.1; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 5'UTR 9..343
XX FT /*tag= a
XX FT 344..1516
XX FT /*tag= b
XX FT /product= "Human glycosyl sulfotransferase-4alpha
XX FT (GST-4alpha)"
XX FT /note= "CDS is specifically claimed as SEQ ID NO: 4
XX FT in claim 6 (page no: 41) of the specification"
XX FT 1517..2134
XX FT 3'UTR /*tag= c
XX
XX WO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 9905-0144694.
XX
XX 13-JUL-2000; 2000US-053828.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX
XX WPI: 2001-138471/14.
XX
XX P-PSDB; AAY72639.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Claim 6; Page 62; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
XX alpha) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation,
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.

SO Sequence 2170 BP; 449 A; 682 C; 612 G; 427 T; 0 other;
Query Match 2.9%; Score 59; DB 22; Length 2170;
Best Local Similarity 100.0%; Pred. No. 3.3e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 678 AGGCTGCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 736
DB 834 AGGCTGCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 892
RESULT 16
ABN89506
ID ABN89506 standard; cDNA: 2544 BP.
XX
AC ABN89506;
XX
DT 05-SEP-2002 (first entry)
XX
XX Human corneal N-acetylglucosamine-6-sulfotransferase cDNA SEQ ID NO:1.
XX
XX Human: N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST.
XX corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
XX opthalmological; chromosome 16q22; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 693..1880
XX FT /*tag= a
XX FT /product= "N-acetylglucosamine-6-sulfotransferase"
XX
XX US2002061562-A1.
XX
XX 23-MAY-2002.
XX
XX 09-AUG-2001; 2001US-0927602.
XX
XX 11-AUG-2000; 2000US-325773P.
XX
XX (FUKU/) FUKUDA M N.
XX (AKAM/) AKAMA T O.
XX
XX Fukuda MN, Akama TO;
XX
XX WPI: 2002-507643/54.
XX
XX P-PSDB; ABB81554.
XX
XX New nucleic acid encoding corneal
XX N-acetylglucosamine-6-sulfotransferase, useful for treatment,
XX monitoring and diagnosis of macular corneal dystrophy
XX
XX Claim 4; Fig 1A-D; 69pp; English.
XX
XX The present sequence encodes human corneal
XX N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse
XX sulfation of keratan sulfate (KS). Also described is a method for
XX monitoring the effect of treatments for macular corneal dystrophy (MCD),
XX and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
XX and has ophthalmological activity. (I) can be used to treat or prevent
XX macular corneal dystrophy types I or II. (I) makes possible treatment
XX of MCD without requiring keratoplasty or keratectomy.
XX
XX Sequence 2544 BP; 460 A; 800 C; 733 G; 551 T; 0 other;
QY 678 AGGCTGCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 736
DB 1180 AGGCTGCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 1238
Query Match 2.9%; Score 59; DB 24; Length 2544;
Best Local Similarity 100.0%; Pred. No. 3.3e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

FT      3'UTR      99662..99968
FT      /*tag= t
XX      MO200106015-A1.
XX      25-JAN-2001.
XX      19-JUL-2000; 2000WO-US19741.
XX      20-JUL-1999; 99US-0144694.
XX      13-JUL-2000; 2000US-0593828.
XX      (REGC ) UNIV CALIFORNIA.
XX      Rosen SD, Lee JK, Hemmerich S;
XX      WPI; 2001-138471/14.
XX      P-PSDB; AAY72639, AAY72640.
XX      New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX      diagnostic and therapeutic agent screening applications
XX      Example 1; Page 62-104; 128pp; English.
XX      The present sequence is human glycosyl sulfotransferase-4 (GST-4) genomic
XX      DNA encoding GST-4alpha and GST-4beta. GST-4 gene is found on
XX      chromosome 10q23.1.
XX      GST is a type 2 membrane protein useful for inhibiting a binding event
XX      between a selectin and a selectin ligand, which comprises contacting the
XX      selectin with a non-sulphated selectin ligand, GST and a small molecular
XX      agent that inhibits the sulphation activity of GST. GST is also useful
XX      in inhibiting a selectin mediated binding event. GST is useful in gene
XX      therapy to treat disorders such as acute or chronic inflammation,
XX      systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX      nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX      glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX      disease, Grave's disease, adrenalitis, hypoparathyroidism, periculous
XX      anaemia, demyelinating diseases, cirrhosis, uicervative colitis,
XX      dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX      syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX      bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX      during transplantation.
XX      Sequence 160552 BP; 40281 A; 37573 C; 38015 G; 44564 T; 119 other;
SQ
Query Match      2.9%; Score 59; DB 22; Length 160552;
Best Local Similarity 100.0%; Pred. No. 2.5e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      678 AGGCGTGGCGCTCTACAGCCAGCGTGTGCTCAAGGAGCGCTTCTTCAACCTGCAG 736
Db      48446 AGGCGTGGCGCTCTACAGCCAGCGTGTGCTCAAGGAGCGCTTCTTCAACCTGCAG 48504

```

```

XX      26-FEB-1999; 99WO-US04316.
XX      20-MAR-1998; 98US-0045284.
XX      12-NOV-1998; 98US-0190911.
XX      (REGC ) UNIV CALIFORNIA.
XX      (SYNT ) SYNTEX USA INC.
XX      Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
XX      WPI; 1999-580442/49.
XX      Human and murine glycosyl sulfotransferase 3 and related
XX      polynucleotides
XX      Example 4; Page 30; 59pp; English.
XX      This sequence encodes the human glycosyl sulfotransferase-3 (GST-3) of
XX      the invention. The nucleic acid sequences, probes and primers derived
XX      from these, proteins and antibodies are useful in detecting homologues.
XX      The sequences, antibodies and methods are useful in the diagnosis and
XX      treatment of diseases associated with selectin binding interactions,
XX      including conditions associated with or resulting from the homing of
XX      leukocytes to sites of inflammation and the normal homing of lymphocytes
XX      to secondary lymph organs.
XX      Sequence 37 BP; 14 A; 9 C; 8 G; 6 T; 0 other;
SQ
Query Match      1.8%; Score 37; DB 20; Length 37;
Best Local Similarity 100.0%; Pred. No. 9.9e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      884 AAACCTCAAGAAGAGGACCAACCTACTATGTATGTC 920
Db      1 AAACCTCAAGAAGAGGACCAACCTACTATGTATGTC 37

```

RESULT 20
ABK70821
ID ABK70821 standard; DNA: 30 BP.
AC ABK70821;
DT 15-JUL-2002 (first entry)
DE Probe for human gene CHST4.
XX Human; sulphuric acid conjugation; ss; probe; CHST1; CHST3;
XX CHST4; CHST5; CST; HNK-1ST; SULFAL; SULFBL; SULFX3; STE; TPST2.
OS Homo sapiens.
XX JP2002085067-A.
XX 26-MAR-2002.
XX 07-SEP-2000; 2000JP-0272229.
XX 07-SEP-2000; 2000JP-0272229.
XX (SAKA) OTSUKA SEIYAKU KOGYO KK.
XX WPI; 2002-378272/41.
XX Determination of enzymes participating in sulphuric acid conjugation in
XX humans, useful for confirmation of safety of investigational drugs,
XX comprises using oligonucleotide probes -
XX Claim 4; Page 9; 13pp; Japanese.
XX The invention relates to classification and quantitative determination of
XX enzymes participating in sulphuric acid conjugation comprising using

CC oligonucleotide probes hybridizing to the following regions: (a) 885-911
CC region of CHST1 gene; (b) 174-197 region of CHST3 gene; (c) 1003-1032
CC region of CHST4 gene; (d) 322-346 region of CHST5 gene; (e) 737-765
CC region of CST gene; (f) 703-732 region of HNK-1ST gene; (g) 299-325
CC region of SULTR1 gene; (h) 358-382 region of SULTR1 gene; (i) 554-582
CC region of SULTR3 gene; (j) 451-478 region of STP gene; and (k) 652-677
CC region of TPST2 gene. Also included are PCR primers for the above
CC genes, kits and methods for determination. The probes, primers and the
CC method are used in the determination of sulphuric acid conjugation for
CC confirmation of the safety of investigational drugs. The present
CC sequence is an oligonucleotide probe for one of the above listed genes.
XX

SO Sequence 30 BP; 5 A; 7 C; 7 G; 11 T; 0 other;

Query Match

Best Local Similarity 1.5%; Score 30; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0096;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1184 CGCTGGTCTTTGGCCATGAAAGCTTTC 1213
DB 1 CGCTGGTCTTTGGCCATGAAAGCTTTC 30

RESULT 21

AAZ20793
ID AAZ20793 standard; DNA; 1926 BP.
XX
AC AAZ20793:
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 coding sequence.
XX
KM Glycosyl sulfotransferase: GST-3; detection; diagnosis; leukocyte homing;
KM selectin binding interaction; inflammation; lymphocyte homing; mouse;
KM secondary lymph organ; ss.
XX
OS Mus sp.
XX
PN WO9449018-A1.
XX
PD 30-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIT CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistrup A, Rosen SD, Tangemann K, Hammerich S;
XX
XX WPI; 1999-580442/49.
DR P-PDB; AAY39919.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
XX polynucleotides
PS Claim 4; Fig 3; 59pp; English.
XX
CC This sequence encodes the mouse glycosyl sulfotransferase-3 (GST-3) of
CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC The sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions,
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.
XX
SQ Sequence 1926 BP; 426 A; 520 C; 484 G; 496 T; 0 other;

Query Match

1.3%; Score 27; DB 20; Length 1926;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 494 GACATGACGCTTTGATGACCTACATG 520
DB 714 GACATGACGCTTTGATGACCTACATG 740

Search completed: February 22, 2003, 20:44:18
Job time : 697 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 20:29:50 ; Search time 95 Seconds

(without alignments)
6559.655 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032 1 ggcgcagagccagagatgcct.....ataagaaaaaaaaaaaaa 2032

Sequence: 1 ggcgcagagccagagatgcct.....ataagaaaaaaaaaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 25

Total number of hits satisfying chosen parameters: 4
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB	ID
1	2032	100.0	2032	4 US-09-045-284A-1
2	2032	100.0	2032	4 US-09-190-911-2
3	37	1.8	37	4 US-09-045-284A-7
4	37	1.8	37	4 US-09-190-911-7

ALIGNMENTS

RESULT 1
US-09-045-284A-1
; Sequence 1, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107051
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-045-284A-1									
Query Match									
Best Local Similarity 100.0%; Score 2032; DB 4; Length 2032;									
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GGCTGAGGCGCAGATGCTCCAGTCTGGGGGAAAATGCTTCCTCATTTGCTTCCAG	60						
DB	1	GGCTGAGGCGCAGATGCTCCAGTCTGGGGGAAAATGCTTCCTCATTTGCTTCCAG	60						
QY	61	CCACCTCAAGCAGTCTCCACCCCTTGAGTCTCAGAGGTAAAGCTTTCTTCA	120						
DB	61	CCACCTCAAGCAGTCTCCACCCCTTGAGTCTCAGAGGTAAAGCTTTCTTCA	120						
QY	121	CAGCTTCTGGGAGAGTCTTCTCAAGCCCTCTTCAAGGTCTTCCACTTCAAGC	180						
DB	121	CAGCTTCTGGGAGAGTCTTCTCAAGCCCTCTTCAAGGTCTTCCACTTCAAGC	180						
QY	181	AATGCTACTGCTTAAAAAATGAAGCTCTGTTCTGCTTTCCAGATGCCATCTT	240						
DB	181	AATGCTACTGCTTAAAAAATGAAGCTCTGTTCTGCTTTCCAGATGCCATCTT	240						
QY	241	GGCTTATTTCCACATGTACAGCACACATCAGCTCCCTGTCTATGAAGGCACAGC	300						
DB	241	GGCTTATTTCCACATGTACAGCACACATCAGCTCCCTGTCTATGAAGGCACAGC	300						
QY	301	CGAGCGCATGCAGTCTGCTGCTCTCTGCTGCGCTCTGCTCTTTTGGGGCA	360						
DB	301	CGAGCGCATGCAGTCTGCTGCTCTCTGCTGCGCTCTGCTCTTTTGGGGCA	360						
QY	361	GCTTTTGGGAGACCCAGATGTTTCTACCTGATGAGGCGCGCGGACGCTGGAT	420						
DB	361	GCTTTTGGGAGACCCAGATGTTTCTACCTGATGAGGCGCGCGGACGCTGGAT	420						
QY	421	GACCTTCAAGCAGACACCGGCTGATGCTGATGCTGCTGCGGATCTGATACGGC	480						
DB	421	GACCTTCAAGCAGACACCGGCTGATGCTGATGCTGCTGCGGATCTGATACGGC	480						
QY	481	GCTTCTTCTGGGAGCAGATGAGCTGTTGATGCTTACATGGAAGCTTCCCGGACACA	540						
DB	481	GCTTCTTCTGGGAGCAGATGAGCTGTTGATGCTTACATGGAAGCTTCCCGGACACA	540						
QY	541	GTCAGCCTCTTTCAGTGGAGAACAGCGCGGCTGTCTGCTACCTGCTGTGACAT	600						
DB	541	GTCAGCCTCTTTCAGTGGAGAACAGCGCGGCTGTCTGCTACCTGCTGTGACAT	600						
QY	601	CATCCCAAAATGAAATCATCCCGGGCTCATGTCAGGCTCTCTGTCAATCAAGCC	660						
DB	601	CATCCCAAAATGAAATCATCCCGGGCTCATGTCAGGCTCTCTGTCAATCAAGCC	660						
QY	661	CTTGGAGTGTGGAGAGGCTGCGCTCTACAGCAGCTGATGTCAGAGAGGTGCG	720						
DB	661	CTTGGAGTGTGGAGAGGCTGCGCTCTACAGCAGCTGATGTCAGAGAGGTGCG	720						
QY	721	CTTCTTCAACTGCAGTCTCTTACCCGCTGCTGAAGACCCCTCTCAACCTGCATAT	780						
DB	721	CTTCTTCAACTGCAGTCTCTTACCCGCTGCTGAAGACCCCTCTCAACCTGCATAT	780						
QY	781	CGTGACCTGGGCGGAGCCCGGGCGGTTCGCTCCGAGAGGCAAAAGGGAGA	840						
DB	781	CGTGACCTGGGCGGAGCCCGGGCGGTTCGCTCCGAGAGGCAAAAGGGAGA	840						
QY	841	TCTCATGATGACAGTCTCATTTGATGGGAGAGATGAGCAAGAACTCAAGAGAGA	900						
DB	841	TCTCATGATGACAGTCTCATTTGATGGGAGAGATGAGCAAGAACTCAAGAGAGA	900						
QY	901	CCAACTTACTATGATGATGAGAGTCTATGTCGAAGCCAGCTGAGATCTCAAGACAT	960						
DB	901	CCAACTTACTATGATGATGAGAGTCTATGTCGAAGCCAGCTGAGATCTCAAGACAT	960						
QY	961	CCAGTCTTGGCCAAAGGCTGTCAGAGAGCTTACTGCTTGTGGCTATGAGAGCTGGC	1020						
DB	961	CCAGTCTTGGCCAAAGGCTGTCAGAGAGCTTACTGCTTGTGGCTATGAGAGCTGGC	1020						

1021 TCGAGCCCTGTGGCCAGAGCTTCCGATGTAATGATGCGGATTTGGATTTGCC 1080
1021 TCGAGCCCTGTGGCCAGAGCTTCCGATGTAATGATGCGGATTTGGATTTGCC 1080
1081 CCAATCTGAGACTGGGTGCAATACATCACCCGAGGCAAGGGATGGGTGACACGCTTT 1140
1081 CCAATCTGAGACTGGGTGCAATACATCACCCGAGGCAAGGGATGGGTGACACGCTTT 1140
1141 CCACACAAATGCGGAGTGGCTTAAATGCTCCAGGCTTGGGCGCTGCTTGGCCCTA 1200
1141 CCACACAAATGCGGAGTGGCTTAAATGCTCCAGGCTTGGGCGCTGCTTGGCCCTA 1200
1201 TGAAGAGTTTCTGACTTGAAGAGGCTTGGCGATGATGATGATTTGCTGGGCTACCG 1260
1201 TGAAGAGTTTCTGACTTGAAGAGGCTTGGCGATGATGATGATTTGCTGGGCTACCG 1260
1261 CCAGCTGAGATCTGAG 1320
1261 CCAGCTGAGATCTGAG 1320
1321 TGTCCCTGAGAGAAATCCACTAGAGGAGTGAAGAGGCTTGTGCCACCTGGTGTAGCC 1380
1321 TGTCCCTGAGAGAAATCCACTAGAGGAGTGAAGAGGCTTGTGCCACCTGGTGTAGCC 1380
1381 TCAATCACTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
1381 TCAATCACTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
1441 TCTGTGGGTATACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
1441 TCTGTGGGTATACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
1501 GTGTCCATGCTTGTGTAGAAAACAGAGCTGGGGAGCTTATGTAGAGAGAGAGAGAGAG 1560
1501 GTGTCCATGCTTGTGTAGAAAACAGAGCTGGGGAGCTTATGTAGAGAGAGAGAGAGAG 1560
1561 CAGTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1561 CAGTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1621 AGACTTGTGGCTGAGAGGCTTATAGACAGACAGATGATGATGATGATGATGATGATGATG 1680
1621 AGACTTGTGGCTGAGAGGCTTATAGACAGACAGATGATGATGATGATGATGATGATGATG 1680
1681 ACCCTGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1681 ACCCTGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1741 ATTTTCCAGAGAGATGCAAAATTCGAGCCCTTGGAGTTCCAGTGTGATTCAGAGAGAA 1800
1741 ATTTTCCAGAGAGATGCAAAATTCGAGCCCTTGGAGTTCCAGTGTGATTCAGAGAGAA 1800
1801 GTGGAGAAAGAGTGGATGCTTACTTATAGCTGAGCAGACAGATGATGATGATGATGATG 1860
1801 GTGGAGAAAGAGTGGATGCTTACTTATAGCTGAGCAGACAGATGATGATGATGATGATG 1860
1861 AATATGAACAAATCTCTGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
1861 AATATGAACAAATCTCTGACAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
1921 TGCATTTGAATATCT 1980
1921 TGCATTTGAATATCT 1980
1981 AAGGTGCACTGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2032
1981 AAGGTGCACTGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2032

RESULT 2
US-09-190-911-2
; Sequence 2, Application US/09190911

Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/09/190,911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045,284
EARLIER FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
TYPE: DNA
ORGANISM: H. sapiens
US-09-190-911-2
Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGTGAGGCGCAGAGATGCTCCAGTCTGGGGGAAATGCTTCTGATTTGCTTCCAG 60
1 GGTGAGGCGCAGAGATGCTCCAGTCTGGGGGAAATGCTTCTGATTTGCTTCCAG 60
61 CCCACCTCAAGCAGATCTCCACCCCTTGTAGTCTGAGCAGATGTTAAAGCTGTACTTCA 120
61 CCCACCTCAAGCAGATCTCCACCCCTTGTAGTCTGAGCAGATGTTAAAGCTGTACTTCA 120
121 CAGCTTCTGAGGAGAGAGTCTTCTCAAGCCCGTCTTCAAGAGTCTTCCACTTAGAGAC 180
121 CAGCTTCTGAGGAGAGAGTCTTCTCAAGCCCGTCTTCAAGAGTCTTCCACTTAGAGAC 180
121 CAGCTTCTGAGGAGAGAGTCTTCTCAAGCCCGTCTTCAAGAGTCTTCCACTTAGAGAC 180
181 AATGCTAGCTTAAATAAATGAAGCTCTGCTGTTTCTGATGATGATGATGATGATGATG 240
181 AATGCTAGCTTAAATAAATGAAGCTCTGCTGTTTCTGATGATGATGATGATGATGATG 240
241 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
241 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
301 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
301 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
361 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
361 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
421 GACCTTCAAGCAG 480
421 GACCTTCAAGCAG 480
481 CCGTCTTCTGTCGAGATGAG 540
481 CCGTCTTCTGTCGAGATGAG 540
541 GTTCAGAGCTTCTTCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 GTTCAGAGCTTCTTCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 CATCCCAAGATGAAAT 660
601 CATCCCAAGATGAAAT 660
661 CTTTGAAGTGTGAG 720
661 CTTTGAAGTGTGAG 720
721 CTTTGAAGTGTGAG 780
721 CTTTGAAGTGTGAG 780

```

Db 721 CTTCTTCAACCTGCACTGCTTACCCTGCTGTAAGAACCCCTCCCTCAACCTCATAT 780
QY 781 CGTGCACCTGTGTCGGGAGACCCCGGCGGTGTCGTTCCGAGAAAGCACAAGGAGAGA 840
Db 781 CGTGCACCTGTGTCGGGAGACCCCGGCGGTGTCGTTCCGAGAAAGCACAAGGAGAGA 840
QY 841 TCTCATGATTGACAGTGCATGATGAGGAGAGATGAGAGAGAACTCAAGAAAGAGAGA 900
Db 841 TCTCATGATTGACAGTGCATGATGAGGAGAGATGAGAGAGAACTCAAGAAAGAGAGA 900
QY 901 CCAGCCCTACTATGTATGACAGTGCATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 CCAGCCCTACTATGTATGACAGTGCATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CCAGTCTCTGCCCCAAGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 CCAGTCTCTGCCCCAAGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 TCGAGCCCCCTGTGCCCCAGACTTCCCGAATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 TCGAGCCCCCTGTGCCCCAGACTTCCCGAATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 CCATCTTACAGACCTGGTGTCATACATACACCCAGGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 CCATCTTACAGACCTGGTGTCATACATACACCCAGGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 CCACACAATGCGAGGAGATGCGCTTAATGTCTCCAGAGGTTGGGCTGTGCTGCTGCTG 1200
Db 1141 CCACACAATGCGAGGAGATGCGCTTAATGTCTCCAGAGGTTGGGCTGTGCTGCTGCTG 1200
QY 1201 TGAAGAAGTTCTGCACTTCAGAAAGCTGTGCGATGCCATGAAATTTGCTGGGCTACCG 1260
Db 1201 TGAAGAAGTTCTGCACTTCAGAAAGCTGTGCGATGCCATGAAATTTGCTGGGCTACCG 1260
QY 1261 CCAGCTCAGATGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 CCAGCTCAGATGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TGTCCCTGAGCAATCCATTAAGAGGTTGAGAGAGCTTTGCTGACACCTGCTGTACAGCC 1380
Db 1321 TGTCCCTGAGCAATCCATTAAGAGGTTGAGAGAGCTTTGCTGACACCTGCTGTACAGCC 1380
QY 1381 TCACTCACTTCTGCAATGCTTCTGAGCCTTGCTACATCTCTGAGCCTTACTACATG 1440
Db 1381 TCACTCACTTCTGCAATGCTTCTGAGCCTTGCTACATCTCTGAGCCTTACTACATG 1440
QY 1441 TCTGTGGGATACACACTGAGTGTGAGTGTGTCACACAGTGTGTCAGAGAGAGAGAGAG 1500
Db 1441 TCTGTGGGATACACACTGAGTGTGAGTGTGTCACACAGTGTGTCAGAGAGAGAGAGAG 1500
QY 1501 GTGTCCATGCTGTGTCTGAGAAACAGAGTGGGAACTTATGTAGAGAGAGAGAGAGAG 1560
Db 1501 GTGTCCATGCTGTGTCTGAGAAACAGAGTGGGAACTTATGTAGAGAGAGAGAGAGAG 1560
QY 1561 CAGTGAACAGAGGATGCTGCTTCTTCTTCTTCTGATCTCTGCTGCTGCTGCTGCTGCTG 1620
Db 1561 CAGTGAACAGAGGATGCTGCTTCTTCTTCTTCTGATCTCTGCTGCTGCTGCTGCTGCTG 1620
QY 1621 AGACTTTGTGGCTGAGAGCCTATTAAAGCAGACAGACAGTATCAGTATGATGATCATTA 1680
Db 1621 AGACTTTGTGGCTGAGAGCCTATTAAAGCAGACAGACAGTATCAGTATGATGATCATTA 1680
QY 1681 ACCCTCCCTGTCACATCTTGGCCATGGGAGATGATCTTTTACCAAGAGAGAGAGAGAG 1740
Db 1681 ACCCTCCCTGTCACATCTTGGCCATGGGAGATGATCTTTTACCAAGAGAGAGAGAGAG 1740
QY 1741 ATTTCACAGAGATGACAAATCTGAGCCCTTGAGGCTTCCAGTGGATGATCAAGAGAGAA 1800
Db 1741 ATTTCACAGAGATGACAAATCTGAGCCCTTGAGGCTTCCAGTGGATGATCAAGAGAGAA 1800
QY 1801 GTGGGAACAAGGTTGATGCTTACTTATGAGGTTGACCATCAGACTATCGGTAATCGA 1860

```

```

Db 1801 GTGGGAACAAGGTTGATGCTTACTTATGAGCTTGACCATCAGACTTGGGTAATCGA 1860
QY 1861 AATATGAACAAATCTCTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 AATATGAACAAATCTCTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 TGCATTTGAATATGACTTCCCTGCTGATTTCCCATCAGATAGAGAGAGAGAGAGAGAGAG 1980
Db 1921 TGCATTTGAATATGACTTCCCTGCTGATTTCCCATCAGATAGAGAGAGAGAGAGAGAG 1980
QY 1981 AAGCTGCCATCTGTTAATCTAAATCCCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2032
Db 1981 AAGCTGCCATCTGTTAATCTAAATCCCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2032

RESULT 3
US-09-045-284A-7
; Sequence 7, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107051
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-045-284A-7

Query Match 1.88; Score 37; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 4,3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTCAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920
Db 1 AAACCTCAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 37

RESULT 4
US-09-190-911-7
; Sequence 7, Application US/09190911
; Patent No. 6365365
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Tangemann, Kirsten
; APPLICANT: Hemmerlich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-10701P
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 37
; TYPE: DNA
; ORGANISM: H. sapiens
US-09-190-911-7

Query Match 1.88; Score 37; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 4,3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTCAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920

```

Db 1 AAAGTCAAGAGAGAGACCAACCTACTATGTGATGC 37

Search completed: February 22, 2003, 22:53:37
Job time : 96 secs

Mon Feb 24 10:56:26 2003

us-09-816-825-1.olig25.rnpb

Page 1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 22:24:31 ; Search time 166 Seconds

(without alignments)
6875.008 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032 1 ggcgcgagcgccagatgcct.....ataaagaaaaaaaaa 2032

Scoring table: OLIGO-MUC
Gapop 60.0 , Gapext 60.0

Searched: 442118 seqs, 280819700 residues

Word size : 25

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2032	100.0	2032	9	US-10-007-262-2
2	2032	100.0	2032	10	US-09-816-825-1
3	505	24.9	505	10	US-09-998-598-2595
4	505	24.9	517	10	US-09-919-580-194
5	286	14.1	389	10	US-09-919-580-264
6	59	2.9	2544	10	US-09-927-602-1
7	59	2.9	48435	10	US-09-927-602-38
8	37	1.8	37	9	US-10-007-262-7
9	37	1.8	37	10	US-09-816-825-7

ALIGNMENTS

RESULT 1
US-10-007-262-2
; Sequence 2, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:

APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OR INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2032
TYPE: DNA
ORGANISM: H. sapiens
US-10-007-262-2

Query Match 100.0%; Score 2032; DB 9; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTGAGGCGCAGATGCTCCAGTCTGGGGGAAATGCTTCCTTCCTCCAG 60
DB 1 GGCCTGAGGCGCAGATGCTCCAGTCTGGGGGAAATGCTTCCTTCCTCCAG 60
QY 61 CCCACTCAAGACAGTCTCCACCCCTTGAATCAGAGTGTAAAGTGTACTTCA 120
DB 61 CCCACTCAAGACAGTCTCCACCCCTTGAATCAGAGTGTAAAGTGTACTTCA 120
QY 121 CAGCTTCTGGAGGAGGAGTCTTCTCAAGCCGCTTCGAAGTCTTCCACTTACAG 180
DB 121 CAGCTTCTGGAGGAGGAGTCTTCTCAAGCCGCTTCGAAGTCTTCCACTTACAG 180
QY 181 AATCTACTGCTCTAAATAAATGAAGCTCTGCTTCTGTTCCAGATGGCACTT 240
DB 181 AATCTACTGCTCTAAATAAATGAAGCTCTGCTTCTGTTCCAGATGGCACTT 240
QY 241 GGCCTTATCTCTTCACATGTACAGCCCAACATCATGCTCCCTGTCTATGAAGCAGCC 300
DB 241 GGCCTTATCTCTTCACATGTACAGCCCAACATCATGCTCCCTGTCTATGAAGCAGCC 300
QY 301 CGAGCGATGACAGCTGCTGCTGCTCTCTCTGCTGAGGCTGCTCTTGTGGGCA 360
DB 301 CGAGCGATGACAGCTGCTGCTGCTCTCTCTGCTGAGGCTGCTCTTGTGGGCA 360
QY 361 GCTTTTGGGAGGAGCCAGATGTTTCTACCTGATGAGCCGCCCTGGCAGCTGTGAT 420
DB 361 GCTTTTGGGAGGAGCCAGATGTTTCTACCTGATGAGCCGCCCTGGCAGCTGTGAT 420
QY 421 GACCTTCAAGAGAGGAGCCAGCTGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 GACCTTCAAGAGAGGAGCCAGCTGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GCTTCTTGTGACATGAGCCGCTTGTGATGCTTACATGGAACCTGCTCCCGGAGACA 540
DB 481 GCTTCTTGTGACATGAGCCGCTTGTGATGCTTACATGGAACCTGCTCCCGGAGACA 540
QY 541 GTCAGGCTCTTCAAGGAGGAGAGACAGCCGCGCTTGTGACAGCTGCTGTGAT 600
DB 541 GTCAGGCTCTTCAAGGAGGAGAGACAGCCGCGCTTGTGACAGCTGCTGTGAT 600
QY 601 CATCCACAGAGATGAATATATCCCGGGGCTCATCTGAGGCTCCTGTGAGTAAACAGCC 660
DB 601 CATCCACAGAGATGAATATATCCCGGGGCTCATCTGAGGCTCCTGTGAGTAAACAGCC 660
QY 661 CTTTGAAGTGTGAGAGAGGCTGCTCCCTCTACAGCAGCTGTGCTCAAGAGGTGGC 720
DB 661 CTTTGAAGTGTGAGAGAGGCTGCTCCCTCTACAGCAGCTGTGCTCAAGAGGTGGC 720
QY 721 CTTTCAACCTGAGTCTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 CTTTCAACCTGAGTCTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Db	481	CGTCTCTTGGGCACATGAGACCGCTCTTGAATGCGCTACATGAAACCGTGTCCCGGAGACA	540
Qy	541	GTCCAGCCTCTTTAGTGGGAGAAACAGCCGGCCCTGTGTCTTCACACTGCTGTGACAT	600
Db	541	GTCCAGCCTCTTTAGTGGGAGAAACAGCCGGCCCTGTGTCTTCACACTGCTGTGACAT	600
Qy	601	CATCCACAATATGAATTCATCCCGCGGGTCACCTAGAGCTCCCTGATGACCAACAGCC	660
Db	601	CATCCACAATATGAATTCATCCCGCGGGTCACCTAGAGCTCCCTGATGACCAACAGCC	660
Qy	661	CTTTGAGTGTGAGAAAGCCCTGCGCTCCACAGCCAGTGTGCTCAGAGAGTGGC	720
Db	661	CTTTGAGTGTGAGAAAGCCCTGCGCTCCACAGCCAGTGTGCTCAGAGAGTGGC	720
Qy	721	CTTTCAACTGAGATCCCTCTACCCGGGTGGAAGAAGCCCTCCCTCAACTGCATAT	780
Db	721	CTTTCTCAACTGAGATCCCTCTACCCGGGTGGAAGAAGCCCTCCCTCAACTGCATAT	780
Qy	781	CGTGCACCTGTGCGGGAGCCCGGGCGGTGTCCCTTCCCGAGAACGACAAAGGAGA	840
Db	781	CGTGCACCTGTGCGGGAGCCCGGGCGGTGTCCCTTCCCGAGAACGACAAAGGAGA	840
Qy	841	TCTCATATTCACATGCGCATTTGTGATGGGGCAGCATGAGCAGAAACTAAGAAAGAGA	900
Db	841	TCTCATATTCACATGCGCATTTGTGATGGGGCAGCATGAGCAGAAACTAAGAAAGAGA	900
Qy	901	CCAAACCCTACTANTGTGATGAGAGTTCATCTCCAAACCAGCTGGAGATATACAAACCAT	960
Db	901	CCAAACCCTACTANTGTGATGAGAGTTCATCTCCAAACCAGCTGGAGATATACAAACCAT	960
Qy	961	CCAGTCTTGGCCCAAGGCCCTGCAAGAGAGCTACCTGTGTGTGCGCTATGAGAGCTGGC	1020
Db	961	CCAGTCTTGGCCCAAGGCCCTGCAAGAGAGCTACCTGTGTGTGCGCTATGAGAGCTGGC	1020
Qy	1021	TCGAGCCCTGTGGCCCAAGCTTCCGAAATGATGATTCGTGGGATTTGGAATTTCTTGGC	1080
Db	1021	TCGAGCCCTGTGGCCCAAGCTTCCGAAATGATGATTCGTGGGATTTGGAATTTCTTGGC	1080
Qy	1081	CCATCTTCAAGACTGGGTGCATATACATCACCCGAGGCAAGGGCATAGGTGACACGCTTT	1140
Db	1081	CCATCTTCAAGACTGGGTGCATATACATCACCCGAGGCAAGGGCATAGGTGACACGCTTT	1140
Qy	1141	CCACACAATATGCCAGGAGATCCCTTAATATGTCGCCAGGCTTGGCCCTGTCTTGGCCCTA	1200
Db	1141	CCACACAATATGCCAGGAGATCCCTTAATATGTCGCCAGGCTTGGCCCTGTCTTGGCCCTA	1200
Qy	1201	TGAAAAGGTTCTCAGACTTCAGAAAGCCGTGTGGCATGTCATGAATTTGCTGGCTACCG	1260
Db	1201	TGAAAAGGTTCTCAGACTTCAGAAAGCCGTGTGGCATGTCATGAATTTGCTGGCTACCG	1260
Qy	1261	CCAGCTAGATTCGAAACAAAGAAAGAAACCTGTGTGATCTTCTGTCTACTGTGAC	1320
Db	1261	CCAGCTAGATTCGAAACAAAGAAAGAAACCTGTGTGATCTTCTGTCTACTGTGAC	1320
Qy	1321	TGTCCCTGAGCAAAATCCACTAAGAGGTTGAGAAGGCTTGTGTCCACCTGGTGTACACC	1380
Db	1321	TGTCCCTGAGCAAAATCCACTAAGAGGTTGAGAAGGCTTGTGTCCACCTGGTGTACACC	1380
Qy	1381	TCAATCACTTTCTCTGAATCTTCTAAGCCTTGCTACATCTCTAGACCTTACTACATG	1440
Db	1381	TCAATCACTTTCTCTGAATCTTCTAAGCCTTGCTACATCTCTAGACCTTACTACATG	1440
Qy	1441	TCTGTGGGATTCACACTGAGATGTGAATTTGTGTCACACAGTGTCAAGAGAAAGGACTTT	1500
Db	1441	TCTGTGGGATTCACACTGAGATGTGAATTTGTGTCACACAGTGTCAAGAGAAAGGACTTT	1500
Qy	1501	GTGTCCATGCTTGTGTACAAAAACAGACTGGGAAACCTTATGTAGACAGCATCCAC	1560
Db	1501	GTGTCCATGCTTGTGTACAAAAACAGACTGGGAAACCTTATGTAGACAGCATCCAC	1560
Qy	1561	CAGTGAAGACGGGATTTGCTCTCTTCTTGTATGATCTCTGCTGTGGCAGACTTCAG	1620
Db	1561	CAGTGAAGACGGGATTTGCTCTCTTCTTGTATGATCTCTGCTGTGGCAGACTTCAG	1620

```

Db 241 GATTTCTCTGTGGGAGACTTCAGAGACTTTGTGGCTGGAGCCCTATTAAAGCAGAC 300
QY 1655 ACAGTATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGGCCCAATGGGAATG 1714
Db 301 ACAGTATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGGCCCAATGGGAATG 360
QY 1715 GATCTTCACCAAGAGCTACACAGATTTTCCACAGAGATGCAATTTGAGCCCTGG 1774
Db 361 GATCTTCACCAAGAGCTACACAGATTTTCCACAGAGATGCAATTTGAGCCCTGG 420
QY 1775 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAGGTGATGCTTACTTATGAGCTT 1834
Db 421 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAGGTGATGCTTACTTATGAGCTT 480
QY 1835 GACCATCACAGCTATGCGTATACAG 1859
Db 481 GACCATCACAGCTATGCGTATACAG 505

RESULT 4
US-09-919-580-194
; Sequence 194, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-194

Query Match          24.9%; Score 505; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 5.1e-205;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 GCGTTTGCTGCCACCTGGGTGTACAGCTCAGTCACTTTCTGTGAATGCTTGTGAGCCTTGC 1414
Db 13 GCGTTTGCTGCCACCTGGGTGTACAGCTCAGTCACTTTCTGTGAATGCTTGTGAGCCTTGC 72
QY 1415 CTACATCTGTGAGCTTAATACATGCTGTGGGTATACACAGTGAATGATGATGCTGC 1474
Db 73 CTACATCTGTGAGCTTAATACATGCTGTGGGTATACACAGTGAATGATGATGCTGC 132
QY 1475 ACAAGTGTCTAGCAGAGAGACTTTGTGTGCATGCTTGTGTAGAAAACAGACTGGGG 1534
Db 133 ACAAGTGTCTAGCAGAGAGACTTTGTGTGCATGCTTGTGTAGAAAACAGACTGGGG 192
QY 1535 AACCTTGTGAGAGACATCCACAGAGTGAAGAGGATGCTTCTTTCTTCTT 1594
Db 193 AACCTTGTGAGAGACATCCACAGAGTGAAGAGGATGCTTCTTTCTTCTT 252
QY 1595 GATCTTCCTGTCTGGGAGACTTCAGAGACTTTGTGGCTGGAGGCTTATTAAGCAGAC 1654
Db 253 GATCTTCCTGTCTGGGAGACTTCAGAGACTTTGTGGCTGGAGGCTTATTAAGCAGAC 312
QY 1655 ACAGTATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGGCCCAATGGGAATG 1714
Db 313 ACAGTATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGGCCCAATGGGAATG 372
QY 1715 GATCTTCACCAAGAGCTACACAGATTTTCCACAGAGATGCAATTTGAGCCCTGG 1774
Db 373 GATCTTCACCAAGAGCTACACAGATTTTCCACAGAGATGCAATTTGAGCCCTGG 432
```

```

QY 1775 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAGGTGATGCTTACTTATGAGCTT 1834
Db 433 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAGGTGATGCTTACTTATGAGCTT 492
QY 1835 GACCATCACAGCTATGCGTATACAG 1859
Db 493 GACCATCACAGCTATGCGTATACAG 517

RESULT 5
US-09-919-580-264
; Sequence 264, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-264

Query Match          14.1%; Score 286; DB 10; Length 389;
Best Local Similarity 100.0%; Pred. No. 5e-112;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1732 CTCACAGCATTTTCCACAGAGATCCAAATTCGAGCCCTTGGAGTCCCAAGTGATCA 1791
Db 13 CTCACAGCATTTTCCACAGAGATCCAAATTCGAGCCCTTGGAGTCCCAAGTGATCA 72
QY 1792 AGGAAGAGAGTGGGAACAAGTGTGATGCTTACTTATGAGCTTGCATCACAGTATGC 1851
Db 73 AGGAAGAGAGTGGGAACAAGTGTGATGCTTACTTATGAGCTTGCATCACAGTATGC 132
QY 1852 GTAATCAGAAATATGAAACAAATCTCTGCACAAAGAGCAAGCTTAAAGTGCAGAGG 1911
Db 133 GTAATCAGAAATATGAAACAAATCTCTGCACAAAGAGCAAGCTTAAAGTGCAGAGG 192
QY 1912 TGCCTGGGCTGATTTGAATATCACTTCCCTGTGCATTTTCCATACATAGAACTT 1971
Db 193 TGCCTGGGCTGATTTGAATATCACTTCCCTGTGCATTTTCCATACATAGAACTT 252
QY 1972 TGACCTGTGAAGCTGCATCTGTATATATAATAATCCCAATAG 2017
Db 253 TGACCTGTGAAGCTGCATCTGTATATATAATAATCCCAATAG 298

RESULT 6
US-09-927-602-1
; Sequence 1, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2544
```

TYPE: DNA
ORGANISM: Homo Sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (693)...(1877)
US-09-927-602-1

Query Match 2.9%; Score 59; DB 10; Length 2544;
Best Local Similarity 100.0%; Pred. No. 1.2e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGCCCTCTACAGCCAGCGTGTCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
DB 1180 AGGCTGCGCCCTCTACAGCCAGCGTGTCTCAAGAGAGTGGCTTCTTCAACCTGCAG 1238

RESULT 7
US-09-927-602-38
Sequence 38, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 48436
TYPE: DNA
ORGANISM: Homo Sapien
US-09-927-602-38

Query Match 2.9%; Score 59; DB 10; Length 48436;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCTGCGCCCTCTACAGCCAGCGTGTCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
DB 48024 AGGCTGCGCCCTCTACAGCCAGCGTGTCTCAAGAGAGTGGCTTCTTCAACCTGCAG 48082

RESULT 8
US-10-007-262-7
Sequence 7, Application US/10007262
Patent No. US20020164748A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 37
TYPE: DNA
ORGANISM: H. sapiens
US-10-007-262-7

Query Match 1.8%; Score 37; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTAAGAAGAGGACCAACCCCTACTANGTATGTC 920
DB 1 AAACCTAAGAAGAGGACCAACCCCTACTANGTATGTC 37

RESULT 9
US-09-816-825-7
Sequence 7, Application US/09816825
Patent No. US20010051370A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 37
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-825-7

Query Match 1.8%; Score 37; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTAAGAAGAGGACCAACCCCTACTANGTATGTC 920
DB 1 AAACCTAAGAAGAGGACCAACCCCTACTANGTATGTC 37

Search completed: February 23, 2003, 00:13:05
Job time : 270 secs

This Page Blank (usps)

Db 361 HVSEGEORMLDLSTWTFEQIH 386

RESULT 2

09R11 ID PRELIMINARY; PRT; 388 AA.

AC 09R11; 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE L-selectin ligand sulfotransferase.

GN CHST4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RA [1]

SEQUENCE FROM N.A.

MDL:LINE=9361934; PubMed=10435561;

RA Hiraka N., Petrylak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,

RA Iwata D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.;

RT "A novel, high endothelial venule-specific sulfotransferase expresses

RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.";

RT Immunol. 11:79-89(1998)

DR EMBL: AF109155; AAD45579.1; -

DR MGD; MGI:1349479; Chst4.

KW Lectin; Selectin; Transferase.

SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884AEE CRC64;

QY 44 HVLVSSWRSGSFVQGLFGQHPDVFYLMPEAMHWMTF 82

Db 43 HVLVSSWRSGSFVQGLFGQHPDVFYLMPEAMHWMTF 81

RESULT 3

09WUE5 ID PRELIMINARY; PRT; 388 AA.

AC 09WUE5; 01-NOV-1999 (TREMblrel. 12, Created)

DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)

DE N-acetylglucosamine 6-O-sulfotransferase (Carbohydrate (Chondroitin

DE 6/Retan) sulfotransferase 4).

GN CHST4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RA [1]

SEQUENCE FROM N.A.

MDL:LINE=9264336; PubMed=1030415;

RA Bistrup A., Bhakta S., Lee J.-K., Below Y.Y., Gunn M.D., Zuo F.-R.,

RA Huang C.-C., Kannagi R., Rosen S.D., Hemmerich S.;

RT "Sulfotransferases of two specificities function in the reconstitution

RT of high endothelial cell ligands for L-selectin.";

RT J. Cell Biol. 145:899-910(1999).

RL [2]

SEQUENCE FROM N.A.

MDL:LINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Konno S., Yamashita T.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Cassavant T.,

RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Marzarelli U., Kommaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AF131236; AAD33016.1; -

DR EMBL; AK009113; BAB26078.1; -

DR MGD; MGI:1349479; Chst4.

KW Transferase.

SEQUENCE 388 AA; 44694 MW; D1E9D7796DF8574D CRC64;

Query Match 10.1%; Score 39; DB 11; Length 388;

Best Local Similarity 100.0%; Pred. No. 1.4e-34;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSFVQGLFGQHPDVFYLMPEAMHWMTF 82

Db 43 HVLVSSWRSGSFVQGLFGQHPDVFYLMPEAMHWMTF 81

RESULT 4

09GZX3 ID PRELIMINARY; PRT; 395 AA.

AC 09GZX3; 01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)

DE N-acetylglucosamine 6-O-sulfotransferase GSN-4beta (Corneal N-

DE acetylglucosamine-6-O-sulfotransferase).

GN GSTA6BTA OR CHST6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

RA [1]

SEQUENCE FROM N.A.

MDL:LINE=20472330; PubMed=11017086;

RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Fujiwara T.,

RA Nakamura T., Dota A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S.,

RA Ozaki K., Kinoshita S., Shimomura Y., Tanigami A., Fukuda M.N.;

RT "Macular corneal dystrophy type I and type II are caused by distinct

RT mutations in a new sulphotransferase gene.";

RL Nat. Genet. 26:237-241(2000).

DR EMBL; AF280086; AAG48244.1; -

DR EMBL; AF219990; AAG36325.1; -

DR EMBL; AF219991; AAG26327.1; -

KW Transferase.

SEQUENCE 395 AA; 44098 MW; 433CA60248A48F7 CRC64;

Query Match 4.9%; Score 19; DB 4; Length 395;

Best Local Similarity 100.0%; Pred. No. 3.5e-12;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSFVQGLF 62

Db 42 HVLVSSWRSGSFVQGLF 60

```

RESULT 5
Q9GZS9 PRELIMINARY; PRT; 395 AA.
AC Q9GZS9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE N-acetylglucosamine 6-O-sulfotransferase.
GN CHST5 OR I-GLCNAC-6-ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
RT Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF176841; AAD56003.1; -
DR EMBL; AF176840; AAD56002.1; -
DR MGI; MGI:1931825; Chst5.
KW SEQUENCE 395 AA; 44537 MW; 3FD71EA3ED383BE CRC64;

Query Match 4.9%; Score 19; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSNRSRSGSFEVQLF 62
DB 42 HVLVLSNRSRSGSFEVQLF 60

RESULT 6
Q9GZS9 PRELIMINARY; PRT; 411 AA.
AC Q9GZS9; Q9GZS9;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Intestinal GlcNAc-6-sulfotransferase (intestinal N-
DE acetylglucosamine-6-O-sulfotransferase) (N-acetylglucosamine
DE 6-O-sulfotransferase).
GN CHST5 OR I-GLCNAC-6-ST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20472330; PubMed=11017086;
RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K.,
RA Nakamura T., Dora A., Kawasaki S., Inoue Y., Maeda S., Yamamoto S.,
RA Fujiwara T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A.,
RA Fukuda M.N.;
RT "Macular corneal dystrophy type I and type II are caused by distinct
RT mutations in a new sulphotransferase gene.";
RL Nat. Genet. 26:237-241(2000).
RN [2]
RP SEQUENCE OF 22-411 FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
RT Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF246718; AAG28023.1; -
DR EMBL; AF249991; AAG26326.1; -
DR EMBL; AF176839; AAD56001.1; -
DR EMBL; AF176838; AAD56000.1; -

```

```

KW Transferrase.
SQ SEQUENCE 411 AA; 46160 MW; 97642D54BE926E06 CRC64;

Query Match 4.9%; Score 19; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSYSHVYKVRFFENLQ 185
DB 186 ACRSYSHVYKVRFFENLQ 204

```

Search completed: February 20, 2003, 13:25:37
Job time : 36 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:23:11 ; Search time 18 Seconds

(without alignments)
630,358 Million cell updates/sec

Title: US-09-816-825-2

Perfect score: 386
Sequence: 1 MLPRKKMLLFLVSQMAIL.....EQRLLDLSTWTVEQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCYUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfilltest.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	386	100.0	386	4	US-09-045-284A-2
2	386	100.0	386	4	US-09-190-911-1

ALIGNMENTS

RESULT 1

US-09-045-284A-2

Sequence 2, Application US/09045284A
Patent No. 6265192
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107051
CURRENT APPLICATION NUMBER: US/09/045, 284A
CURRENT FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 100.0%; Score 386; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKKMLLFLVSQMAILALFFHMTSHNISTSMKRAQPERMVLVLSWRSSSVGQ	60
DB	1	MLPRKKMLLFLVSQMAILALFFHMTSHNISTSMKRAQPERMVLVLSWRSSSVGQ	60
QY	61	LFQHPDVFYLMPEPAHWMTFFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGRRQ	120
DB	61	LFQHPDVFYLMPEPAHWMTFFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGRRQ	120
QY	121	SSLFQWENSRLCSPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
DB	121	SSLFQWENSRLCSPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
QY	181	FFNQSLYPLKDPNLNHLVLRDPAVFRSRERTKGDLMTDSRTVMGHEQKLKED	240
DB	181	FFNQSLYPLKDPNLNHLVLRDPAVFRSRERTKGDLMTDSRTVMGHEQKLKED	240
QY	241	QPYVVMQYICQSLERIKTISLPKALQERYLVRYEDLAPAPVAGTSRMTEYFGLFPL	300
DB	241	QPYVVMQYICQSLERIKTISLPKALQERYLVRYEDLAPAPVAGTSRMTEYFGLFPL	300
QY	301	HLQTMVHNTKSGKMGDAFHTNARDALNVSQARWMSLPYEKYSRLKACGDANLLGYR	360
DB	301	HLQTMVHNTKSGKMGDAFHTNARDALNVSQARWMSLPYEKYSRLKACGDANLLGYR	360
QY	361	HVRSQEQRLNLLDLSTWTVEQIH 386	
DB	361	HVRSQEQRLNLLDLSTWTVEQIH 386	

RESULT 2

US-09-190-911-1

Sequence 1, Application US/09190911
Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107051
CURRENT APPLICATION NUMBER: US/09/190, 911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045, 284
EARLIER FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-09-190-911-1

Query Match 100.0%; Score 386; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKKMLLFLVSQMAILALFFHMTSHNISTSMKRAQPERMVLVLSWRSSSVGQ	60
DB	1	MLPRKKMLLFLVSQMAILALFFHMTSHNISTSMKRAQPERMVLVLSWRSSSVGQ	60
QY	61	LFQHPDVFYLMPEPAHWMTFFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGRRQ	120
DB	61	LFQHPDVFYLMPEPAHWMTFFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGRRQ	120
QY	121	SSLFQWENSRLCSPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
DB	121	SSLFQWENSRLCSPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
QY	181	FFNQSLYPLKDPNLNHLVLRDPAVFRSRERTKGDLMTDSRTVMGHEQKLKED	240

Db 181 FFNLSLYPLKDPSSLNLHIVLVRDPAVFRSRERTKGDLMDSRIVMGQHEOKLKED 240
QY 241 QPYVWGVICQSQLEIYKTIQSLPKALOERYLLVRYEDLARAPVQTSRWYEEVGLFLP 300
Db 241 QPYVWGVICQSQLEIYKTIQSLPKALOERYLLVRYEDLARAPVQTSRWYEEVGLFLP 300
QY 301 HLOTWVHNITRGKMGDGHAFHTNARDALNYSQAMRWSLPEYKYSRLQKACGDAMNLIQYR 360
Db 301 HLOTWVHNITRGKMGDGHAFHTNARDALNYSQAMRWSLPEYKYSRLQKACGDAMNLIQYR 360
QY 361 HVRSEQEQRLNLLDLSTWTVPEQIH 386
Db 361 HVRSEQEQRLNLLDLSTWTVPEQIH 386

Search completed: February 20, 2003, 13:26:29
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:22:51 ; Search time 21 Seconds
(without alignments)
1767.042 Million cell updates/sec

Title: US-09-816-825-2

Perfect score: 386

Sequence: 1 MLDPKKMKLLFLVSQMAIL.....EQRNLLDLSTWTPQIH 386

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

No matches found

Search completed: February 20, 2003, 13:26:05
Job time: 21 secs

This page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:19:55 ; Search time 14 Seconds
(Without alignments)
1143.562 Million cell updates/sec

Title: US-09-816-825-2

Perfect score: 386

Sequence: 1 MLPRKMKLLFLVSGMAIL.....EQRNLDDLSTWVPEQIH 386

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
---------------	----------------	-------	--------	----	-------------

No matches found

Search completed: February 20, 2003, 13:24:55
Job time : 14 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:25:41 ; Search time 13 Seconds
(without alignments)
758.605 Million cell updates/sec

Title: US-09-816-825-2

Sequence: 1 MLPRKMKLLFLVSQMAIL.....EQRNLLDLSTWVPEQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 140259 seqs, 25548876 residues

Word size : 15

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	386	100.0	386	9	US-10-007-262-1
2	386	100.0	386	9	US-09-816-825-2
3	60	15.5	171	10	US-09-927-602-8
4	19	4.9	169	10	US-09-927-602-6
5	19	4.9	169	10	US-09-927-602-7
6	19	4.9	390	10	US-09-927-602-4
7	19	4.9	395	10	US-09-927-602-2
8	19	4.9	395	10	US-09-927-602-3
9	19	4.9	418	10	US-09-927-602-5

ALIGNMENTS

RESULT 1
US-10-007-262-1
; Sequence 1, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:

APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 386; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKMKLLFLVSQMAILALFFHMTSHNITSSLMKAQPERMHVLYSSWRSGSSFFVGQ	60
DB	1	MLPRKMKLLFLVSQMAILALFFHMTSHNITSSLMKAQPERMHVLYSSWRSGSSFFVGQ	60
QY	61	LFQHPDVFYLMPEPAWHMTFKOSTAMLMHAYRDLIRAFILCDMSVFDAYPEGRQ	120
DB	61	LFQHPDVFYLMPEPAWHMTFKOSTAMLMHAYRDLIRAFILCDMSVFDAYPEGRQ	120
QY	121	SSLFOWNSRALCAPACDIIPDEIIPRAHCRLLCSQPEVEVEKACRSYSHVLEKEVR	180
DB	121	SSLFOWNSRALCAPACDIIPDEIIPRAHCRLLCSQPEVEVEKACRSYSHVLEKEVR	180
QY	181	FFNQSLYPLPKDSLNHTYHLYVDRPAVRSRERKGMIDSRVYMGQHEQKLEKED	240
DB	181	FFNQSLYPLPKDSLNHTYHLYVDRPAVRSRERKGMIDSRVYMGQHEQKLEKED	240
QY	241	QPYVYMOYICQSQLEIKTIQSLPKALQERYLTVREYEDLARAPVQTSRMVEFGLEFLP	300
DB	241	QPYVYMOYICQSQLEIKTIQSLPKALQERYLTVREYEDLARAPVQTSRMVEFGLEFLP	300
QY	301	HLQTVHNTITGKMGDHAFTNARDALNVSQANRWSLPYKYSRLQACGDANNLLGYR	360
DB	301	HLQTVHNTITGKMGDHAFTNARDALNVSQANRWSLPYKYSRLQACGDANNLLGYR	360
QY	361	HVRSEQEQRNLLDLSTWVPEQIH 386	
DB	361	HVRSEQEQRNLLDLSTWVPEQIH 386	

RESULT 2
US-09-816-825-2
; Sequence 2, Application US/09816825
; Patent No. US20010051370A1
; GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-825-2

```

Query Match          100.0%; Score 386; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLPKKKLLFLVSOAMALALFEHHYNNISLSMKAOPEHHVVLSSWSSGSSFFGQ 60
   |||
DB 1 MLPLPKKKLLFLVSOAMALALFEHHYNNISLSMKAKAPERHHVVLSSWSSGSSFFGQ 60

QY 61 LFGQHPDVEFLMEPAHVMWTFKOSTAMMLHMAVRDLIAVFLCDMSVFDAYMEGPGRQ 120
   |||
DB 61 LFGQHPDVEFLMEPAHVMWTFKOSTAMMLHMAVRDLIAVFLCDMSVFDAYMEGPGRQ 120

QY 121 SSLFOWENSRALCSAPACDIIPDEIIPRAHCHLLCSQGFVEVEKACRSYSHVYLKEVR 180
   |||
DB 121 SSLFOWENSRALCSAPACDIIPDEIIPRAHCHLLCSQGFVEVEKACRSYSHVYLKEVR 180

QY 181 FFNIQSLYPLPKDPSLNHLIHLVDRPRAVFRSERTKGDMLDSRTVNGQHBOKLKKED 240
   |||
DB 181 FFNIQSLYPLPKDPSLNHLIHLVDRPRAVFRSERTKGDMLDSRTVNGQHBOKLKKED 240

QY 241 QPYVVMQVTCOSQLEIYKTIQS.PKALOERYLLIYREDLARAPVQTSRYEFVGLFPLP 300
   |||
DB 241 QPYVVMQVTCOSQLEIYKTIQS.PKALOERYLLIYREDLARAPVQTSRYEFVGLFPLP 300

QY 301 HLQTVHNITRGKMGDHAETNARDALNVSAWMSLPEYKYSRLQKACGDAMNLGGR 360
   |||
DB 301 HLQTVHNITRGKMGDHAETNARDALNVSAWMSLPEYKYSRLQKACGDAMNLGGR 360

QY 361 HVSEQRRLDLILSTWVPEQIH 386
   |||
DB 361 HVSEQRRLDLILSTWVPEQIH 386

RESULT 3
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8

Query Match          15.5%; Score 60; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 7,3e-50;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 KACRSYSHVYLKEVRFNLSQSLYPLKDPSLNHLIHLVDRPRAVFRSERTKGDMLDS 225
   |||
DB 53 KACRSYSHVYLKEVRFNLSQSLYPLKDPSLNHLIHLVDRPRAVFRSERTKGDMLDS 112

RESULT 4
US-09-927-602-6
; Sequence 6, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-6

```

```

; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-6

Query Match          4.9%; Score 19; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSWRSGSSFFVGOLF 62
   |||
DB 10 HVLVLSWRSGSSFFVGOLF 28

RESULT 5
US-09-927-602-7
; Sequence 7, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-7

Query Match          4.9%; Score 19; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSYSHVYLKEVRFNLSQ 185
   |||
DB 51 ACRSYSHVYLKEVRFNLSQ 69

RESULT 6
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-4

```


US-09-927-602-4

Query Match 4.9%; Score 19; DB 10; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSSHWLKEVREFNLO 185

DB 165 ACRSSHWLKEVREFNLO 183

RESULT 7

US-09-927-602-2
; Sequence 2, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-2

Query Match 4.9%; Score 19; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSFFVGOLF 62

DB 42 HVLVSSWRSGSFFVGOLF 60

RESULT 8

US-09-927-602-3
; Sequence 3, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: VARIANT
; LOCATION: (1)...(395)
; OTHER INFORMATION: Xaa = any amino acid
US-09-927-602-3

Query Match 4.9%; Score 19; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSFFVGOLF 62
DB 42 HVLVSSWRSGSFFVGOLF 60

RESULT 9

US-09-927-602-5
; Sequence 5, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-927-602-5

Query Match 4.9%; Score 19; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSFFVGOLF 62
DB 66 HVLVSSWRSGSFFVGOLF 84

Search completed: February 20, 2003, 13:30:26
Job time : 13 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:11 ; Search time 51.3714 Seconds
(without alignments)
75.222 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78
Perfect score: 166
Sequence: 1 SWRSGSFVQGLFGQHPDYFLMEPAMHV 29

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	166	100.0	171 23	ABB81560
2	166	100.0	380 23	AAU11274
3	166	100.0	386 23	AAV39918
4	166	100.0	386 21	AAV92919
5	166	100.0	388 20	AAV33039
6	166	100.0	388 20	AAV39919
7	160	96.4	169 23	ABB81558
8	160	96.4	395 22	AAV72638
9	160	96.4	395 22	AAV72640
10	160	96.4	395 23	ABB81554

11	160	96.4	395 23	ABB81555	Consensus N-acetyl
12	160	96.4	395 23	AAE15438	Human drug metabol
13	160	96.4	395 23	AAU11275	Murine intestinal-
14	160	96.4	418 23	ABB81557	Murine intestinal N
15	157	94.6	169 23	ABB81559	Human intestinal N
16	157	94.6	390 22	AAV72639	Human glycosyl sul
17	157	94.6	390 23	ABB81556	Human glycosyl sul
18	157	94.6	418 21	AAAB1947	Human intestinal N
19	157	94.6	169 23	ABB81561	Human ORFX ORP1711
20	124	74.7	483 20	AAV31656	Human N-acetylgluc
21	124	74.7	484 20	AAV31657	Human N-acetylgluc
22	124	74.7	530 22	AAV95367	Human N-acetylgluc
23	124	74.7	531 22	AAV95414	Human protein sequ
24	116	69.9	179 23	ABB81562	Lung small cell ca
25	116	69.9	411 19	AAAB1100	Human Keratan sulf
26	101	60.8	174 23	ABB81563	Keratan sulphate 6
27	101	60.8	458 18	AAV06480	Human chondroitin-
28	101	60.8	479 19	AAV06480	Chick chondroitin
29	71	42.8	315 22	ABB64513	Glycosaminoglycan
30	53.5	32.2	372 18	AAV08432	Drosophila melanog
31	50	30.1	344 22	AAV08432	Human interleukin-
32	50	30.1	791 22	AAE22917	C glutamicum prote
33	50	30.1	4097 22	ABB57815	Human transporter
34	49.5	29.8	465 22	ABB82580	Drosophila melanog
35	49.5	29.8	265 22	AAV3389	Human caspase recr
36	49	29.5	265 23	AAV97518	Human protein sequ
37	49	29.5	265 23	AAV75501	Novel human protei
38	49	29.5	265 23	AAV75504	Human s acyl fatty
39	49	29.5	318 22	AAV3499	S acyl fatty acid
40	49	29.5	318 22	AAV3499	Human protein sequ
41	48.5	29.2	931 20	AAV75505	S acyl fatty acid
42	48.5	29.2	1250 21	AAV72357	Group B Streptococ
43	48.5	29.2	1252 22	AAV91279	Group B Streptococ
44	48.5	29.2	1252 23	ABP26711	Streptococcus poly
45	48.5	29.2	1252 23	ABP26711	Streptococcus poly
46	48	28.9	240 23	ABB9146	Listeria monocytog
47	48	28.9	288 22	ABB33040	Escherichia coli p
48	48	28.9	460 22	AAU82710	Pseudomonas aerugi
49	48	28.9	753 23	AAU82710	Amio acid sequenc
50	48	28.9	817 22	ABB66979	Drosophila melanog
51	48	28.9	1083 22	ABB61710	Drosophila melanog
52	48	28.9	3606 22	ABB62595	Drosophila melanog
53	47	28.3	222 22	AAU62676	Protonibacterium
54	47	28.3	346 23	ABP26749	Streptococcus poly
55	47	28.3	349 23	ABB33502	Streptococcus poly
56	47	28.3	397 21	AAV33211	Lactococcus lactis
57	47	28.3	701 22	AAU36457	Human ORFX ORP2975
58	47	28.3	925 22	AAV36457	Pseudomonas aerugi
59	47	28.3	1068 23	ABB98243	C glutamicum prote
60	46.5	28.0	99 23	AAE16011	Human CGDP S60 ID
61	46.5	28.0	109 23	ABB60076	Human ga_9504136 p
62	46.5	28.0	261 23	AAE16026	Human DITRP polype
63	46.5	28.0	261 23	AAE16021	Human caspase-12 i
64	46.5	28.0	288 23	AAE16033	Human caspase-12 i
65	46.5	28.0	288 23	AAE16033	Human RIK-4 protei
66	46.5	28.0	340 23	AAE16019	Human RIK-5 protei
67	46.5	28.0	341 22	AAE16019	Human caspase-12 i
68	46.5	28.0	341 22	AAE16005	Human caspase recr
69	46.5	28.0	372 23	AAE16029	Human caspase-12 i
70	46.5	28.0	373 23	AAE16030	Human caspase-12 i
71	46.5	28.0	373 23	AAE16032	Human caspase-12 p
72	46.5	28.0	376 17	AAV44448	Card gene product.
73	46.5	28.0	378 23	AAE19167	Human protease, PR
74	46.5	28.0	378 23	AAU72861	Human aspartyl pro
75	46.5	28.0	397 22	ABG11652	Novel human diago
76	46.5	28.0	669 23	ABP29904	Streptococcus poly
77	46.5	28.0	669 23	ABP29904	Streptococcus poly
78	46.5	28.0	669 23	ABP29904	Streptococcus poly
79	46.5	28.0	1635 22	ABB59277	Hericicidally activ
80	46	27.7	302 22	AAV69992	Drosophila melanog
81	46	27.7	453 22	ABG13564	C glutamicum prote
82	46	27.7	557 23	ABB55154	Novel human diago
83	46	27.7	786 22	ABG04819	Lactococcus lactis
					Novel human diago

84	45.5	27.4	481	21	AAV58629	Protein regulating
85	45.5	27.4	708	21	AAAB40833	Human ORFX ORF617
86	45.5	27.4	709	23	ABBP41264	Human ovarian anti
87	45.5	27.4	844	21	ABAB56597	Synechocystis sp p
88	45.5	27.4	4547	22	ABAB59051	Drosophila melanog
89	45	27.1	82	22	ABG008485	Novel human diagno
90	45	27.1	99	22	ABBA43514	Peptide #11020 enc
91	45	27.1	99	22	ABBA43514	Peptide #8473 enc
92	45	27.1	99	22	AAAM64451	Human brain expres
93	45	27.1	99	22	AAAM77257	Human bone marrow
94	45	27.1	99	22	AAAM21193	Peptide #7627 enc
95	45	27.1	99	22	AAAM37408	Peptide #11445 enc
96	45	27.1	99	22	ABG46273	Human peptide epco
97	45	27.1	145	22	AAAM92877	Human digestive sy
98	45	27.1	156	22	AAAB76790	Corynebacterium gl
99	45	27.1	160	22	AAU421440	Propionibacterium
100	45	27.1	165	23	ABPA0486	Staphylococcus epi
101	45	27.1	202	22	AAAM79674	Human protein SRQ
102	45	27.1	213	22	AAAM78690	Human protein SRQ
103	45	27.1	221	22	AAAB29791	Human steroidogeni
104	45	27.1	236	23	ABP25590	Streptococcus poly
105	45	27.1	239	21	AAAB53294	Human colon cancer
106	45	27.1	239	22	AAU03124	Streptococcus pyog
107	45	27.1	239	22	AAAG75034	Human colon cancer
108	45	27.1	239	23	ABP25591	Streptococcus poly
109	45	27.1	303	23	ABAB89991	Human polypeptide
110	45	27.1	346	20	AAAY35973	Extended human sec
111	45	27.1	355	22	ABG28149	Novel human diagno
112	45	27.1	361	22	ABG27289	Novel human diagno
113	45	27.1	376	22	AAAG2435	C glutamicum prote
114	45	27.1	417	22	AAAM52361	T26F17.15 protein.
115	45	27.1	417	23	ABAB91091	Herbicidally activ
116	45	27.1	418	23	ABAB56693	R. norvegicus RAS
117	45	27.1	643	15	AAAB60020	Fliorectin. Homo
118	45	27.1	644	16	AAAB82244	Human fibrinogen A
119	45	27.1	742	22	AAAB45904	Human tumor-associ
120	45	27.1	791	20	AAAY41040	Human lung tumor a
121	45	27.1	791	21	AAAB11325	Human lung cancer-
122	45	27.1	791	23	ABAB74957	Human lung tumor
123	45	27.1	847	21	AAAY82891	Alphab subunit of
124	45	27.1	881	23	ABAB75049	Human lung tumor
125	45	27.1	920	23	ABAB75001	Human lung tumor
126	45	27.1	942	21	AAAY70459	Human membrane cha
127	45	27.1	943	20	AAAB74823	Human ICAC-2 prot
128	45	27.1	943	20	AAAY41036	Human lung tumor a
129	45	27.1	943	21	AAAB11321	Human lung cancer-
130	45	27.1	943	23	AAAB86031	Human calcium-acti
131	45	27.1	943	23	ABAB74953	Human lung tumor
132	45	27.1	943	23	AAAB50344	Human calcium acti
133	45	27.1	1279	23	ABAB53787	Lactococcus lactis
134	45	27.1	1291	16	AAAB75201	Tyrosine phosphata
135	45	27.1	1336	15	AAAB60019	Tissue-binding hyb
136	45	27.1	1336	21	AAAG58404	Arabidopsis thalia
137	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
138	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
139	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
140	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
141	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
142	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
143	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
144	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
145	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
146	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
147	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
148	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
149	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
150	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
151	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
152	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
153	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
154	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
155	45	27.1	1336	21	AAAG58403	Arabidopsis thalia
156	45	27.1	1336	21	AAAG58403	Arabidopsis thalia

230	42.5	25.6	210	23	ABP28214	303	42	25.3	526	22	ABB60414	Drosophila melanog
231	42.5	25.6	229	21	AAE66114	304	42	25.3	529	23	AAU74387	Breast tumour-spec
232	42.5	25.6	229	21	AAE65237	305	42	25.3	531	20	AAU25937	Breast tumour-spec
233	42.5	25.6	304	23	ABB84806	306	42	25.3	532	20	AAU25982	Mistletoe lectin I
234	42.5	25.6	348	22	AAAB93467	307	42	25.3	545	22	AAU40860	Mistletoe lectin I
235	42.5	25.6	453	22	AAAB8601	308	42	25.3	544	18	AAU10021	Human polypeptide
236	42.5	25.6	514	21	AAAG54762	309	42	25.3	564	20	AAU90127	Prepro mistletoe I
237	42.5	25.6	590	23	ABP29967	310	42	25.3	616	20	AAU42287	Mistletoe lectin p
238	42.5	25.6	604	23	ABP25463	311	42	25.3	616	20	AAU27399	Human anaphorless p
239	42.5	25.6	614	22	AAU67485	312	42	25.3	621	21	AAU19285	Human P450 reducta
240	42.5	25.6	657	22	ABE68446	313	42	25.3	621	21	AAU19285	A polypeptide with
241	42.5	25.6	823	21	AAU75441	314	42	25.3	656	21	AAU28629	Human B1A91 antiq
242	42.5	25.6	871	21	AAAG49323	315	42	25.3	656	21	AAU82018	Human immunogenic
243	42.5	25.6	875	21	AAAG49322	316	42	25.3	656	22	AAU65977	B305D isoform C sp
244	42.5	25.6	877	21	AAAG49321	317	42	25.3	656	22	AAU69778	Human prostate-cp
245	42.5	25.6	960	23	ABE93642	318	42	25.3	656	22	AAU99018	Human prostate-cp
246	42.5	25.6	2304	23	ABE60219	319	42	25.3	656	22	AAU74816	Human prostate-cp
247	42	25.3	22	16	AAAB81866	320	42	25.3	656	23	AAU95328	Prostate tumour an
248	42	25.3	61	21	AAAG22574	321	42	25.3	656	23	AAU74378	Human B305D splice
249	42	25.3	89	22	AAAG13619	322	42	25.3	671	21	AAU82019	Human B1A91 antiq
250	42	25.3	95	23	ABE80785	323	42	25.3	671	21	AAU65978	Human immunogenic
251	42	25.3	99	22	AAAB84777	324	42	25.3	671	21	AAU69779	B305D isoform C sp
252	42	25.3	108	21	AAAG02246	325	42	25.3	671	22	AAU69779	Human prostate-cp
253	42	25.3	111	21	AAU04889	326	42	25.3	671	22	AAU01134	Human prostate-cp
254	42	25.3	135	22	AAU59289	327	42	25.3	671	22	AAU99019	Human prostate-cp
255	42	25.3	151	22	ABG21145	328	42	25.3	671	22	AAU74817	Human prostate-cp
256	42	25.3	163	23	ABG60948	329	42	25.3	671	23	ABE95239	Prostate tumour an
257	42	25.3	170	20	AAAB89055	330	42	25.3	671	23	AAU74379	Human B305D splice
258	42	25.3	170	20	ABE81236	331	42	25.3	676	20	AAU42286	Breast tumour-spec
259	42	25.3	217	23	ABG60947	332	42	25.3	676	20	AAU42286	Human P450 reducta
260	42	25.3	225	22	AAAB96159	333	42	25.3	676	20	AAU27398	Human P450 reducta
261	42	25.3	250	20	AAAG4294	334	42	25.3	876	22	AAU17538	Human P450 reducta
262	42	25.3	251	23	ABG60930	335	42	25.3	876	22	AAU37030	Staphylococcus aur
263	42	25.3	252	19	AAAB64659	336	42	25.3	1019	23	AAU92447	Herbicidally activ
264	42	25.3	252	19	AAAB64651	337	42	25.3	1354	22	ABE60329	Drosophila melanog
265	42	25.3	253	18	AAAB10022	338	42	25.3	1719	21	AAU82017	Human immunogenic
266	42	25.3	253	20	AAAG90125	339	42	25.3	1719	22	AAU69777	Human prostate-cp
267	42	25.3	254	20	AAU25980	340	42	25.3	1719	22	AAU01132	Human prostate-cp
268	42	25.3	254	20	AAU25983	341	42	25.3	1719	22	AAU99017	Human prostate-cp
269	42	25.3	258	22	ABE68035	342	42	25.3	1719	22	AAU74815	Human prostate-cp
270	42	25.3	260	22	AAAB1774	343	42	25.3	1841	21	AAU07562	Human B305D splice
271	42	25.3	268	22	AAU78652	344	42	25.3	1841	21	AAU07562	Protein encoded by
272	42	25.3	268	22	AAU39537	345	42	25.3	53	23	ABE3798	Human GRF2771 prot
273	42	25.3	273	21	ABE97278	346	42	25.3	56	22	AAAB80384	Human haematologic
274	42	25.3	328	21	AAAG26632	347	42	25.3	90	22	AAU22597	Novel human colon
275	42	25.3	328	22	AAU79636	348	42	25.3	90	22	AAU22597	Human digestive sy
276	42	25.3	328	22	AAU79636	349	42	25.3	105	22	ABE27682	Peptide #358 encod
277	42	25.3	335	21	AAAB2990	350	42	25.3	105	22	ABE27682	Peptide #333 encod
278	42	25.3	335	21	AAAB24423	351	42	25.3	105	22	ABE27682	Peptide #333 encod
279	42	25.3	335	21	AAU93376	352	42	25.3	105	22	AAU93376	Human brain expres
280	42	25.3	335	21	AAU94899	353	42	25.3	105	22	AAU66038	Human bone marrow
281	42	25.3	335	21	AAAB8486	354	42	25.3	105	22	AAU13907	Peptide #341 encod
282	42	25.3	335	22	AAAB6125	355	42	25.3	105	22	AAU26313	Peptide #350 encod
283	42	25.3	335	22	ABE95517	356	42	25.3	105	22	ABE35686	Peptide #332 encod
284	42	25.3	335	23	ABE84911	357	42	25.3	105	23	AAU35686	Human polypeptide
285	42	25.3	335	23	AAU83676	358	42	25.3	127	22	AAU07803	Human polypeptide
286	42	25.3	378	21	AAAB43212	359	42	25.3	182	17	ABE37997	Staphylococcus epi
287	42	25.3	384	21	AAAB28628	360	42	25.3	198	13	AAAB85214	Blowfly larvae PM9
288	42	25.3	384	21	AAAG65976	361	42	25.3	246	22	AAAB84036	Human immunohaema
289	42	25.3	384	23	ABE60999	362	42	25.3	334	22	AAAB6366	Putative p. abyssal
290	42	25.3	384	23	ABE61001	363	42	25.3	407	22	AAU40808	Human polypeptide
291	42	25.3	384	23	AAU74377	364	42	25.3	437	20	AAAB8342	Human polypeptide
292	42	25.3	384	23	AAU74377	365	42	25.3	446	22	AAU39022	Salmonella enteric
293	42	25.3	391	22	ABE16370	366	42	25.3	446	22	AAU62275	Human polypeptide
294	42	25.3	435	20	AAU42288	367	42	25.3	455	22	AAE03202	Human gene 17-enco
295	42	25.3	435	20	AAU42288	368	42	25.3	455	23	ABE64398	Human albumin fusi
296	42	25.3	490	22	ABG04895	369	42	25.3	480	17	AAU85213	Human secreted pro
297	42	25.3	492	22	ABE66209	370	42	25.3	508	22	AAU664726	Blowfly larvae PM9
298	42	25.3	492	22	AAU71914	371	42	25.3	553	22	ABE60817	Murine transcriptp1
299	42	25.3	501	22	AAU43557	372	42	25.3	582	22	AAU62276	Drosophila melanog
300	42	25.3	501	22	AAU43620	373	42	25.3	597	21	AAU30765	Human gene 17-enco
301	42	25.3	501	22	AAU73855	374	42	25.3	645	21	AAU45573	Arabidopsis thalia
302	42	25.3	517	22	AAU65240	375	42	25.3	667	23	ABE63572	Human albumin fusi

376	41.5	25.0	683	22	AAB62172	Human gene 17-enco
377	41.5	25.0	700	22	AAB95135	Human protein sequ
378	41.5	25.0	702	20	AAV14969	Amino acid sequenc
379	41.5	25.0	702	22	AAB83278	Human FATP3 SEQ ID
380	41.5	25.0	730	20	AAV41699	Human PRO703 prote
381	41.5	25.0	730	21	AAB44255	Human PRO703 (UNQ3
382	41.5	25.0	730	21	AAB24054	Human PRO703 prote
383	41.5	25.0	730	22	AAB83279	Human FATP3 SEQ ID
384	41.5	25.0	730	22	AAB60388	Human fatty acid t
385	41.5	25.0	762	21	AAG30764	Arabidopsis thalia
386	41.5	25.0	764	21	AAG30763	Arabidopsis thalia
387	41.5	25.0	764	23	AAB92076	Herbicidally activ
388	41.5	25.0	810	21	AAG45572	Arabidopsis thalia
389	41.5	25.0	812	21	AAV45571	Arabidopsis thalia
390	41.5	25.0	1387	21	AAV95441	Ceonothadictis ele
391	41.5	25.0	5385	22	AAB66487	Drosophila melanog
392	41.5	25.0	5496	22	ABB67161	Drosophila melanog
393	41.5	25.0	8805	22	ABB67112	Drosophila melanog
394	41	24.7	40	9	AA80076	Cyomegalovirus (C
395	41	24.7	45	22	AAB15718	Human nervous syst
396	41	24.7	66	22	AAW65430	Human Immune/haema
397	41	24.7	80	17	AAW01186	Serine protease Pf
398	41	24.7	80	22	AAW50583	Flea serine protea
399	41	24.7	87	22	ABW15699	Human nervous syst
400	41	24.7	96	21	AAW42133	Human ORFX ORF1897
401	41	24.7	109	22	ABB66484	Drosophila melanog
402	41	24.7	113	22	AAW44053	Proiontacterium
403	41	24.7	135	21	AAW29878	Human secreted pro
404	41	24.7	140	22	ABG11570	Novel human diagn
405	41	24.7	140	22	ABG18668	Novel human diagn
406	41	24.7	140	22	ABG18974	Novel human diagn
407	41	24.7	141	21	AAW04689	Arabidopsis thalia
408	41	24.7	149	15	AAW52603	Allinase fragment
409	41	24.7	150	18	AAW26749	S. carnosus nitrat
410	41	24.7	151	22	ABB66220	Drosophila melanog
411	41	24.7	162	21	AAW04688	Arabidopsis thalia
412	41	24.7	167	21	AAW14252	Arabidopsis thalia
413	41	24.7	168	21	AAW14251	Arabidopsis thalia
414	41	24.7	176	22	AAW37397	Human colon cancer
415	41	24.7	183	21	AAW04687	Arabidopsis thalia
416	41	24.7	183	21	AAW53375	Arabidopsis thalia
417	41	24.7	189	15	AAW65911	Recombinant apoaeq
418	41	24.7	189	17	AAW05057	Aequorin D1175 mut
419	41	24.7	189	17	AAW05058	Aequorin D1185 mut
420	41	24.7	189	17	AAW05059	Aequorin G122A mut
421	41	24.7	189	19	AAW79061	Apoaequorin polype
422	41	24.7	189	19	AAW62538	Wild type apoaequ
423	41	24.7	189	19	AAW44736	Apo-aequorin gener
424	41	24.7	189	22	AAW51167	Aequorin photoprot
425	41	24.7	189	22	AAW51171	Recombinant apoaeq
426	41	24.7	189	23	AAW47869	Aequorea luminesce
427	41	24.7	189	23	AAW13374	Aequorea victoria
428	41	24.7	189	23	AAW13378	Aequorea victoria
429	41	24.7	190	21	AAW14250	Arabidopsis thalia
430	41	24.7	196	7	AAW60273	Sequence encoding
431	41	24.7	196	7	AAW60795	Aequorin protein.
432	41	24.7	196	11	AAW04822	Aequorin protein.
433	41	24.7	196	11	AAW65910	Recombinant apoaeq
434	41	24.7	196	15	AAW65912	Recombinant aequo
435	41	24.7	196	15	AAW65913	Recombinant aequo
436	41	24.7	196	15	AAW65914	Recombinant aequo
437	41	24.7	196	16	AAW77875	Apoaequorin. Aequ
438	41	24.7	196	17	AAW05054	Apoaequorin D124S
439	41	24.7	196	17	AAW05055	Apoaequorin E135S
440	41	24.7	196	17	AAW05056	Apoaequorin G129A
441	41	24.7	196	19	AAW59921	Apoaequorin polype
442	41	24.7	196	19	AAW62540	Wild type apoaequ
443	41	24.7	196	19	AAW57996	Apoaequorin #1. A
444	41	24.7	196	19	AAW57997	Apoaequorin #2. A
445	41	24.7	196	19	AAW44737	Wild type apo-aequ
446	41	24.7	196	22	AAW51163	Apoaequorin-encodi
447	41	24.7	196	22	AAW51164	Recombinant aequo
448	41	24.7	196	22	AAW51165	Recombinant aequo
449	41	24.7	196	22	AAW51168	Aequorin mutant pr
450	41	24.7	196	22	AAW51169	Aequorin mutant pr
451	41	24.7	196	22	AAW51170	Aequorin mutant pr
452	41	24.7	196	23	AAW13370	Aequorea victoria
453	41	24.7	196	23	AAW13371	Aequorea victoria
454	41	24.7	196	23	AAW13372	Aequorea victoria
455	41	24.7	196	23	AAW13373	Aequorea victoria
456	41	24.7	196	23	AAW13374	Aequorea victoria
457	41	24.7	197	23	AAW13375	Aequorea victoria
458	41	24.7	197	23	AAW13376	Aequorea victoria
459	41	24.7	210	19	AAW44738	Chimeric apo-aequo
460	41	24.7	210	19	AAW44739	Mutant apo-aequo
461	41	24.7	211	13	AAW22228	Truncated human pr
462	41	24.7	211	13	AAW24273	Truncated human pr
463	41	24.7	214	21	AAW53374	Arabidopsis thalia
464	41	24.7	234	21	AAW53375	Arabidopsis thalia
465	41	24.7	250	22	AAW18450	Human endocrine po
466	41	24.7	254	21	AAW05719	Arabidopsis thalia
467	41	24.7	254	21	AAW42656	Arabidopsis thalia
468	41	24.7	256	22	AAW62508	Flea serine protea
469	41	24.7	267	21	AAW42655	Arabidopsis thalia
470	41	24.7	268	21	AAW05718	Arabidopsis thalia
471	41	24.7	277	21	AAW04206	Novel human diagn
472	41	24.7	279	21	AAW88627	H. influenzae dime
473	41	24.7	298	21	AAW13073	Arabidopsis thalia
474	41	24.7	298	21	AAW42600	Arabidopsis thalia
475	41	24.7	303	21	AAW329725	Arabidopsis thalia
476	41	24.7	348	22	AAW63659	Drosophila melanog
477	41	24.7	349	18	AAW08101	G-protein coupled
478	41	24.7	349	20	AAW17749	Human HIBB51 rece
479	41	24.7	349	21	AAW96921	Soluble human prol
480	41	24.7	349	21	AAW95527	Human prolactin re
481	41	24.7	349	21	AAW90621	Human G protein-co
482	41	24.7	349	21	AAW90655	Human mutant G-pro
483	41	24.7	350	21	AAW70341	Human G protein-co
484	41	24.7	356	21	AAW29724	Arabidopsis thalia
485	41	24.7	356	21	AAW29723	Arabidopsis thalia
486	41	24.7	365	21	AAW13072	Arabidopsis thalia
487	41	24.7	365	21	AAW42599	Arabidopsis thalia
488	41	24.7	376	21	AAW13071	Arabidopsis thalia
489	41	24.7	377	21	AAW53367	Arabidopsis thalia
490	41	24.7	378	23	AAW20067	Lactobacillus rhom
491	41	24.7	393	19	AAW55091	Streptococcus pneu
492	41	24.7	393	23	AAW54585	S. pneumoniae Sp03
493	41	24.7	398	22	AAW00864	Human bone marrow
494	41	24.7	416	21	AAW42598	Herbicidally activ
495	41	24.7	416	21	AAW92139	Arabidopsis thalia
496	41	24.7	423	21	AAW81577	Streptococcus pneu
497	41	24.7	429	19	AAW40381	S. glaucescens abo
498	41	24.7	432	23	AAW08630	GFP fusion protein
499	41	24.7	441	21	AAW44890	Corn chlorismate sy
500	41	24.7	441	23	AAW08631	GFP fusion protein
501	41	24.7	450	23	AAW08632	GFP fusion protein
502	41	24.7	468	23	AAW08633	GFP fusion protein
503	41	24.7	473	22	AAW91788	C glutamicum prote
504	41	24.7	475	23	AAW27899	Streptococcus poly
505	41	24.7	477	23	AAW08634	GFP fusion protein
506	41	24.7	483	21	AAW53366	Arabidopsis thalia
507	41	24.7	487	22	AAW00628	Novel human protei
508	41	24.7	492	22	AAW52501	Escherichia coli p
509	41	24.7	511	21	AAW343615	Human cancer assoc
510	41	24.7	511	22	AAW78811	Human protein SEQ
511	41	24.7	511	22	AAW98239	Human polypeptide
512	41	24.7	520	22	AAW41025	Human polypeptide
513	41	24.7	525	22	AAW11786	Human secreted pro
514	41	24.7	525	22	AAW79795	Human protein SEQ
515	41	24.7	530	22	AAW34512	E. coli cellular p
516	41	24.7	531	22	AAW38406	Salmonella typhi c
517	41	24.7	539	22	AAW06630	Novel human protei
518	41	24.7	551	23	AAW79450	Galactose-recognt
519	41	24.7	569	22	AAW57025	Proiontacterium
520	41	24.7	586	22	AAW00629	Novel human protei
521	41	24.7	589	21	AAW53365	Arabidopsis thalia
522	41	24.7	622	12	AAW10795	Human prolactin re

522	41	24.7	622	23	AAU99354	Human prolactin re
523	41	24.7	622	23	AAU99355	Human prolactin re
524	41	24.7	645	22	AAW78857	Human protein SEQ
525	41	24.7	671	23	ABBS3934	Lactococcus lactis
526	41	24.7	674	22	ABG16936	Novel human diagno
527	41	24.7	686	22	AA862155	Novel human protei
528	41	24.7	702	21	AA840701	Human ORFX ORF465
529	41	24.7	702	22	AAW79025	Human protein SEQ
530	41	24.7	702	22	AA862154	Novel human protei
531	41	24.7	702	22	AA846595	Human testis-speci
532	41	24.7	702	22	AAU72887	Human aspartyl pro
533	41	24.7	715	22	AAU00670	Human TANGO 229 po
534	41	24.7	723	22	AA862153	Novel human protei
535	41	24.7	739	22	AA862152	Novel human protei
536	41	24.7	758	22	ABG28652	Novel human diagno
537	41	24.7	826	22	ABG21979	Novel human diagno
538	41	24.7	906	22	AAU34974	Enterococcus faeca
539	41	24.7	906	23	AB808635	GFP fusion protein
540	41	24.7	939	22	ABG28653	Novel human diagno
541	41	24.7	1092	22	AAW79841	Human protein SEQ
542	41	24.7	1130	22	ABG12663	Novel human diagno
543	41	24.7	1131	22	ABG07281	Novel human diagno
544	41	24.7	1213	22	ABG20896	Novel human diagno
545	41	24.7	1550	22	ABG09297	Novel human diagno
546	41	24.7	2045	22	AB861383	Drosophila melanog
547	40.5	24.4	69	21	AA833745	Arabidopsis thalia
548	40.5	24.4	96	22	AB816535	Human nervous syst
549	40.5	24.4	98	22	AB814724	Human nervous syst
550	40.5	24.4	109	22	AAO01024	Human polypeptide
551	40.5	24.4	171	20	AAW86357	Partial human DNAX
552	40.5	24.4	171	23	AAE16108	Human DNAX Toll 11
553	40.5	24.4	171	23	AAE16113	Human DNAX Toll 11
554	40.5	24.4	231	20	AAV73924	Human prostate tum
555	40.5	24.4	272	21	AA833740	Arabidopsis thalia
556	40.5	24.4	300	22	AB833355	Novel human diagno
557	40.5	24.4	364	22	AB813513	Novel human diagno
558	40.5	24.4	426	21	AA845483	Arabidopsis thalia
559	40.5	24.4	432	21	AA845482	Arabidopsis thalia
560	40.5	24.4	435	21	AA845481	Arabidopsis thalia
561	40.5	24.4	513	22	AB864791	Drosophila melanog
562	40.5	24.4	545	22	AA865649	Novel protein kina
563	40.5	24.4	547	22	AB814111	Novel human diagno
564	40.5	24.4	590	23	ABP29650	Streptococcus poly
565	40.5	24.4	592	23	ABP25462	Streptococcus poly
566	40.5	24.4	624	22	AA890535	C glutamicum prote
567	40.5	24.4	628	22	AA876784	Corynebacterium gl
568	40.5	24.4	628	22	AA876842	Corynebacterium gl
569	40.5	24.4	666	23	ABP28725	Streptococcus poly
570	40.5	24.4	729	23	AB881039	Herbicidially activ
571	40.5	24.4	789	22	AB86162	Drosophila melanog
572	40.5	24.4	846	22	AB870852	Drosophila melanog
573	40.5	24.4	1032	23	AAE16107	Human DNAX Toll 11
574	40.5	24.4	1033	22	ABG23220	Novel human diagno
575	40.5	24.4	4630	18	AAW19629	Streptomyces venez
576	40.5	24.4	4630	21	AAV77177	S. venezuelae vep
577	40.5	24.4	5087	22	AAU10700	Amino acid sequenc
578	40.5	24.4	5588	22	AAV79301	SpnE a polyketide
579	40.5	24.4	5588	20	AA870969	S. spinosa protein
580	40.5	24.1	62	22	AAW08451	T cell epitope CD-
581	40	24.1	17	18	AAU59460	Protonbacterium
582	40	24.1	86	22	AAU53666	Protonbacterium
583	40	24.1	93	22	AA833763	Human immune/naema
584	40	24.1	117	22	AAU53339	Protonbacterium
585	40	24.1	123	20	AAV29655	Human endostatin a
586	40	24.1	126	20	AAV14415	Human secreted pro
587	40	24.1	132	21	AA818010	Pinus radiata pine
588	40	24.1	142	21	AA800840	Human secreted pro
589	40	24.1	145	22	AAW78571	Human protein SEQ
590	40	24.1	145	22	AA873683	Human oxidoreducta
591	40	24.1	156	22	AAW79555	Human protein SEQ
592	40	24.1	160	22	ABG22261	Novel human diagno
593	40	24.1	185	22	AB859421	Drosophila melanog
594	40	24.1	188	21	AA832721	Eucalyptus grandis
595	40	24.1	188	22	AAU48219	Propionibacterium
596	40	24.1	193	21	AAV74784	Neisseria gonorrhoe
597	40	24.1	200	23	AB853912	Lactococcus lactis
598	40	24.1	202	22	AB870894	Drosophila melanog
599	40	24.1	216	15	AA852029	Protein with Oxa
600	40	24.1	224	22	ABG01708	Novel human diagno
601	40	24.1	232	22	AB864845	Drosophila melanog
602	40	24.1	234	22	AA882743	S. epidermidis ope
603	40	24.1	246	23	ABP38291	Staphylococcus epi
604	40	24.1	253	22	AAE20273	Human lung speci
605	40	24.1	253	23	ABG23654	Novel human diagno
606	40	24.1	258	21	AA851378	Arabidopsis thalia
607	40	24.1	260	22	AA866333	Arabidopsis thalia
608	40	24.1	263	23	AAU75502	P acnes linoleate
609	40	24.1	277	22	AAU31582	Rat S-acyl fatty a
610	40	24.1	293	22	AB867961	Novel human secret
611	40	24.1	296	22	AAU62038	Drosophila melanog
612	40	24.1	314	21	AA851377	Propionibacterium
613	40	24.1	314	22	AA893826	Arabidopsis thalia
614	40	24.1	315	22	AA828466	Human protein sequ
615	40	24.1	325	21	AA851376	Sequence encoded b
616	40	24.1	336	23	ABP41646	Arabidopsis thalia
617	40	24.1	339	22	AA879468	Human ovarian anti
618	40	24.1	347	17	AA885334	Corynebacterium gl
619	40	24.1	347	18	AAW15094	Human prostate/col
620	40	24.1	348	17	AA885333	Form 2 of prostate
621	40	24.1	348	18	AAW15093	Human prostate/col
622	40	24.1	348	23	ABG61850	Form 1 of prostate
623	40	24.1	348	23	ABG61832	Prostate cancer-as
624	40	24.1	359	20	AAV06628	Prostate cancer-as
625	40	24.1	359	21	AAV84309	C. elegans tyrosyl
626	40	24.1	377	22	AA890119	A tyrosylprotein s
627	40	24.1	385	15	AA845751	C glutamicum prote
628	40	24.1	385	18	AAW06598	Erynia amylovora
629	40	24.1	385	22	AB858253	Hypersensitive tes
630	40	24.1	403	19	AAW75863	Drosophila melanog
631	40	24.1	403	19	AAW62455	Erynia amylovora
632	40	24.1	403	19	AAW61114	Erynia amylovora
633	40	24.1	403	21	AAW87659	Hypersensitive tes
634	40	24.1	403	21	AAV71093	A hypersensitive r
635	40	24.1	403	21	AA848654	Erynia amylovora
636	40	24.1	403	22	AAE06710	A hypersensitive r
637	40	24.1	403	23	AB809225	Erynia amylovora
638	40	24.1	403	23	AAE18295	Erynia amylovora
639	40	24.1	403	23	AAE16447	E. amylovora hyper
640	40	24.1	404	21	AA838988	Arabidopsis thalia
641	40	24.1	423	20	AAW73657	M. tuberculosis thalia
642	40	24.1	423	20	AAW73767	M. tuberculosis an
643	40	24.1	423	22	AAU08228	M. tuberculosis an
644	40	24.1	423	22	AAU01891	Polypeptide encode
645	40	24.1	423	22	AAE17579	M. tuberculosis an
646	40	24.1	437	21	AA838997	Mycobacterium spec
647	40	24.1	446	19	AAW37905	Arabidopsis thalia
648	40	24.1	454	8	AA870164	Mycobacterium tube
649	40	24.1	460	21	AA838996	Extracellular alka
650	40	24.1	460	21	AA852373	Arabidopsis thalia
651	40	24.1	461	19	AAW76000	Arabidopsis thalia
652	40	24.1	468	23	AB848511	Ornithine decarbox
653	40	24.1	473	20	AAW88332	Listeria monocytog
654	40	24.1	474	22	AAW87963	Salmonella enteric
655	40	24.1	485	22	AB864596	Amino acid sequenc
656	40	24.1	489	21	AAI69408	Drosophila melanog
657	40	24.1	500	21	AAI69409	Mutine insulin-res
658	40	24.1	500	22	AA868187	Human insulin-res
659	40	24.1	513	22	AA891018	Centrosomal protei
660	40	24.1	513	22	AA879563	C glutamicum prote
661	40	24.1	517	21	AAI69879	Corynebacterium gl
662	40	24.1	528	22	AA879409	B. lactofermentum
663	40	24.1	529	23	AB866731	Corynebacterium gl
664	40	24.1	535	19	AAW57570	Human novel polype
665	40	24.1	548	22	AA890101	Homo sapiens B258
666	40	24.1	548	22	AA879408	C glutamicum prote
667	40	24.1	561	21	AAI69411	Corynebacterium gl
						Human insulin-resp

668	40	24.1	589	21	AAG45984	Aradidopsis thalia	74.1	39.5	23.8	268	22	AAB47310	CBF-1. Physcomitr
669	40	24.1	596	22	AAV72641	Human glycosyl sul	74.2	39.5	23.8	280	21	AAG53828	Aradidopsis thalia
670	40	24.1	620	22	ABW2157	Drosophila melanog	74.3	39.5	23.8	323	19	AAW7829	Rat melanocortin-3
671	40	24.1	651	22	ABW58403	Drosophila melanog	74.4	39.5	23.8	323	20	AAW79276	Rat melanocortin-3
672	40	24.1	689	12	AAK14585	Human ras proximat	74.5	39.5	23.8	323	21	AAV94427	Rhesus monkey meta
673	40	24.1	694	21	AAG31953	Aradidopsis thalia	74.6	39.5	23.8	323	22	AAV50828	Rat melanocortin-3
674	40	24.1	710	20	AAV32066	Myobacterium tube	74.7	39.5	23.8	323	22	AAE60968	Rat melanocortin-3
675	40	24.1	710	20	AAE17588	Myobacterium spec	74.8	39.5	23.8	323	22	AAE04830	Rat melanocortin r
676	40	24.1	723	21	AAV74595	Antigenic fusion p	74.9	39.5	23.8	323	21	AAE05095	Mus musculus melan
677	40	24.1	728	21	AAV69410	Murine insulin-res	75.0	39.5	23.8	333	21	AAG20277	Aradidopsis thalia
678	40	24.1	732	21	AAG45983	Aradidopsis thalia	75.1	39.5	23.8	333	21	AAK43039	Aradidopsis thalia
679	40	24.1	733	23	ABP30271	Streptococcus poly	75.2	39.5	23.8	333	22	AAG33827	Aradidopsis thalia
680	40	24.1	743	21	AAG31952	Aradidopsis thalia	75.3	39.5	23.8	333	22	AAE89669	Rice photosensitiv
681	40	24.1	747	23	ABP25975	Streptococcus poly	75.4	39.5	23.8	333	23	AAO15492	Beta vulgaris case
682	40	24.1	760	21	AAG45982	Aradidopsis thalia	75.5	39.5	23.8	333	23	AAE52836	Physcomitrella pat
683	40	24.1	777	22	ABG09295	Novel human diago	75.6	39.5	23.8	336	23	AAE24232	Streptomyces venez
684	40	24.1	788	20	AAV00937	A. lithotrophicus	75.7	39.5	23.8	337	21	AAE18651	Amino acid sequenc
685	40	24.1	788	23	AAE22114	Archaeoglobus lith	75.8	39.5	23.8	337	21	AAV77134	S. venezuelae deso
686	40	24.1	794	23	ABP35694	Fungal ZPC protein	75.9	39.5	23.8	337	21	AAV70999	S. venezuelae deso
687	40	24.1	804	23	ABP91419	Herbicidally activ	76.0	39.5	23.8	337	21	AAV57215	ORF 14 encoded dnd
688	40	24.1	856	20	AAV32064	Myobacterium tube	76.1	39.5	23.8	381	22	AAU54765	Propionibacterium
689	40	24.1	859	23	AAV74593	Antigenic fusion p	76.2	39.5	23.8	404	22	AAG64723	Human transcriptio
690	40	24.1	876	19	AAW69349	Hepatitase prote	76.3	39.5	23.8	404	23	ABW57218	Mouse ischaemic co
691	40	24.1	912	20	AAV16588	A protein that int	76.4	39.5	23.8	415	19	AAW69228	Human lysosomal si
692	40	24.1	912	20	AAV07379	Presenilin binding	76.5	39.5	23.8	430	19	AAK30784	Aradidopsis thalia
693	40	24.1	941	18	AAW31559	Xenopus frog prote	76.6	39.5	23.8	432	20	AAV34623	Chlamydia pneumoni
694	40	24.1	952	22	ABW59168	Drosophila melanog	76.7	39.5	23.8	432	21	AAK31788	Aradidopsis thalia
695	40	24.1	985	14	AAK42995	Glycosyltransferas	76.8	39.5	23.8	468	21	AAK20783	Aradidopsis thalia
696	40	24.1	985	14	AAK42214	Aspergillus niger	76.9	39.5	23.8	472	22	AAU54831	Propionibacterium
697	40	24.1	1006	22	ABG06272	Novel human diago	77.0	39.5	23.8	500	22	AAK81223	Myobacterium tube
698	40	24.1	1006	22	ABG14517	Novel human diago	77.1	39.5	23.8	513	22	AAU14351	Human novel protei
699	40	24.1	1006	22	ABG17623	Novel human diago	77.2	39.5	23.8	513	22	AAK83573	Human protein sequ
700	40	24.1	1050	23	ABW06245	Bacillus circulans	77.3	39.5	23.8	513	22	AAK94606	Human protein sequ
701	40	24.1	1222	22	AAV72642	Human glycosyl sul	77.4	39.5	23.8	524	21	AAK64612	Aradidopsis thalia
702	40	24.1	1374	19	AAW72225	HSV-2 strain SB5 C	77.5	39.5	23.8	533	23	ABK47948	Listeria monocytog
703	40	24.1	1374	19	AAW69753	Herpes simplex vir	77.6	39.5	23.8	534	21	AAK64611	Aradidopsis thalia
704	40	24.1	1384	19	AAW72224	HSV-2 strain SB5 C	77.7	39.5	23.8	541	20	AAV36832	Chlamydia trachoma
705	40	24.1	1396	19	AAW72117	HSV-2 strain SB5 C	77.8	39.5	23.8	587	23	ABW54271	Lactococcus lactis
706	40	24.1	1396	19	AAW72039	HSV-2 strain SB5 C	77.9	39.5	23.8	632	21	ABW6916	Aradidopsis thalia
707	40	24.1	1396	23	ABW3023	Herbicidally activ	78.0	39.5	23.8	681	21	AAK64610	Nitric oxide reduct
708	40	24.1	1449	20	AAV38624	Neisseria meningit	78.1	39.5	23.8	736	22	ABG24252	Aradidopsis thalia
709	40	24.1	1491	22	AAU34072	Staphylococcus aur	78.2	39.5	23.8	789	22	ABK11347	Novel human diago
710	40	24.1	1502	22	AAU36569	Staphylococcus aur	78.3	39.5	23.8	859	16	AAK2886	Human M-phase phos
711	40	24.1	2008	18	AAW2016	Utrrophin truncated	78.4	39.5	23.8	859	19	AAW31227	Human leucine zipp
712	40	24.1	2013	22	AAW67964	Amino acid sequenc	78.5	39.5	23.8	877	19	AAW59040	Barley alpha-gluc
713	40	24.1	2154	20	AAW81639	Mouse elf-1 protei	78.6	39.5	23.8	888	23	ABW57049	Mouse ischaemic co
714	40	24.1	2464	22	AAW78919	Human protein SMO	78.7	39.5	23.8	1067	21	AAK66928	Aradidopsis thalia
715	40	24.1	3433	18	AAW22017	Utrrophin. Homo sa	78.8	39.5	23.8	1079	21	AAK66928	Aradidopsis thalia
716	39.5	23.8	51	17	AAW03193	Human ORFX protein	78.9	39.5	23.8	1086	21	AAG46928	Aradidopsis thalia
717	39.5	23.8	52	17	AAW03198	Serine protease Pf	79.0	39.5	23.8	1095	23	ABW51562	Herbicidally activ
718	39.5	23.8	52	22	ABW5095	Human pancreatic c	79.1	39.5	23.8	1167	23	AAU37982	Streptococcus poly
719	39.5	23.8	61	22	ABW10731	Human digestive sy	79.2	39.5	23.8	1169	22	AAU37855	Streptococcus pneu
720	39.5	23.8	61	22	AAW93029	Human ORFX protein	79.3	39.5	23.8	1169	22	AAU38100	Novel human diago
721	39.5	23.8	62	23	ABP10919	Zea mays protein f	79.4	39.5	23.8	1223	22	ABK19958	Aradidopsis thalia
722	39.5	23.8	86	21	AAK18522	Zea mays protein f	79.5	39.5	23.8	1223	22	ABW55036	Drosophila melanog
723	39.5	23.8	124	21	AAK18521	Benzene dioxigenas	79.6	39.5	23.8	1566	22	ABG04042	Novel human diago
724	39.5	23.8	135	14	AAK32087	Novel human diago	79.7	39.5	23.8	3782	21	AAV77179	S. venezuelae deso
725	39.5	23.8	142	22	ABG12065	Propionibacterium	79.8	39.5	23.8	3782	23	AAE24228	Streptomyces venez
726	39.5	23.8	146	22	AAU50155	Human ORFX ORF86	79.9	39.5	23.8	4572	19	AAE2845	A. mediterranei ri
727	39.5	23.8	159	21	AAK41112	Human transport pr	80.0	39.5	23.5	17	18	AAW08452	T cell epitope CD-
728	39.5	23.8	159	23	ABP34929	Aradidopsis thalia	80.1	39.5	23.5	24	20	AAV25803	Human secreted pro
729	39.5	23.8	176	21	AAK20279	Aradidopsis thalia	80.2	39.5	23.5	32	19	AAW74851	Human secreted pro
730	39.5	23.8	176	21	AAK3041	Aradidopsis thalia	80.3	39.5	23.5	57	22	AAO06280	Human polypeptide
731	39.5	23.8	186	21	AAG20278	Aradidopsis thalia	80.4	39.5	23.5	63	22	ABK44204	Peptide #1170 enc
732	39.5	23.8	186	21	AAG3040	Aradidopsis thalia	80.5	39.5	23.5	63	22	ABK27082	Protein #9081 enco
733	39.5	23.8	186	21	AAG53829	Aradidopsis thalia	80.6	39.5	23.5	63	22	AAK65235	Human brain expres
734	39.5	23.8	192	20	AAW88733	Secreted protein e	80.7	39.5	23.5	63	22	AAW77938	Human bone marrow
735	39.5	23.8	192	22	ABW50507	Human secreted pro	80.8	39.5	23.5	63	22	AAW21837	Peptide #8267 enco
736	39.5	23.8	213	21	AAG49357	Aradidopsis thalia	80.9	39.5	23.5	63	22	AAK38153	Human peptide enco
737	39.5	23.8	220	21	AAG49356	Human colon cancer	81.0	39.5	23.5	66	22	ABG46959	Human immune/haema
738	39.5	23.8	252	17	AAW53418	Serine protease Pf	81.1	39.5	23.5	79	20	AAV73995	Human prostate tum
739	39.5	23.8	252	17	AAW01210	Human secreted pro	81.2	39.5	23.5	80	21	AAW28066	Human secreted pro
740	39.5	23.8	252	22	ABW50607		81.3	39.5	23.5	80	21	AAW28066	

814	39	23.5	81	22	AAU54037	Proprionbacterium
815	39	23.5	90	22	AAU47665	Proprionbacterium
816	39	23.5	100	22	ABG26255	Novel human diagno
817	39	23.5	106	22	AAAB2701	Human immune/haema
818	39	23.5	111	21	AAAB56402	Human prostate can
819	39	23.5	115	22	ABG24876	Novel human diagno
820	39	23.5	118	21	AAAG03122	Human secreted pro
821	39	23.5	123	23	ABP05536	Human ORFX protein
822	39	23.5	126	21	AAAG01064	Human secreted pro
823	39	23.5	128	20	AAU73871	Human prostate tum
824	39	23.5	129	23	ABP02468	Human ORFX protein
825	39	23.5	133	20	AAU13128	Human secreted pro
826	39	23.5	137	22	AAAB0886	Protein encoded by
827	39	23.5	140	23	AAAB7274	Relaxin/IGF/Insuli
828	39	23.5	148	23	AAO21537	Human novel secret
829	39	23.5	154	22	AAU16203	Human novel secret
830	39	23.5	154	22	AAU16616	Human novel secret
831	39	23.5	155	20	AAU28684	Human yb37_1 secre
832	39	23.5	155	21	AAAB5326	Eucalyptus grandis
833	39	23.5	155	23	AAU39040	Human secreted pro
834	39	23.5	155	23	ABAB55749	Human polypeptide
835	39	23.5	156	22	AAU43052	Proprionbacterium
836	39	23.5	156	23	AAAB18616	Human osteopontin-
837	39	23.5	160	21	AAAG38832	Arabidopsis thalia
838	39	23.5	166	21	AAAG38832	Arabidopsis thalia
839	39	23.5	170	23	AAAB18615	Human osteopontin-
840	39	23.5	171	18	AAAB55641	H. pylori ORF 05ep
841	39	23.5	171	19	AAAB98398	H. pylori GHP 118
842	39	23.5	171	21	AAU74785	Neisseria meningit
843	39	23.5	175	22	ABG30076	Novel human diagno
844	39	23.5	178	22	ABBA3725	Peptide #11231 enc
845	39	23.5	178	22	ABAB26671	Protein #8670 enco
846	39	23.5	178	22	AAAB46687	Human brain expres
847	39	23.5	178	22	AAAB21388	Peptide #7822 enco
848	39	23.5	178	22	AAAB37623	Peptide #11660 enc
849	39	23.5	184	22	AAAB3431	Human protein sequ
850	39	23.5	184	23	ABP27840	Streptococcus poly
851	39	23.5	184	23	ABP29837	Streptococcus poly
852	39	23.5	193	21	AAU74786	Neisseria meningit
853	39	23.5	198	20	AAU38371	Protein phosphatas
854	39	23.5	198	20	AAU28623	Human secreted pro
855	39	23.5	198	21	AAAB19602	Human dual-specific
856	39	23.5	198	21	AAAB19620	Dual-specificity p
857	39	23.5	198	21	AAAB19621	Human secreted pro
858	39	23.5	198	21	AAU73394	Human bone marrow
859	39	23.5	214	22	AAU71924	Ferredoxin-HPLV-II
860	39	23.5	215	18	AAW26616	Arabidopsis thalia
861	39	23.5	220	21	AAAG17171	Arabidopsis thalia
862	39	23.5	220	21	AAAG39595	Corynebacterium q1
863	39	23.5	221	22	AAAB79804	S. epidermidis ope
864	39	23.5	221	22	AAAB79804	S. epidermidis ope
865	39	23.5	221	22	ABP39559	Staphylococcus epi
866	39	23.5	226	22	ABG28521	Novel human diagno
867	39	23.5	244	18	AAW06460	Serratia marcescen
868	39	23.5	244	19	AAW1033	Acid phosphatase e
869	39	23.5	245	12	ABAB61836	Drosophila melanog
870	39	23.5	250	21	AAAG17170	Arabidopsis thalia
871	39	23.5	250	21	AAAG39594	Arabidopsis thalia
872	39	23.5	253	21	AAAG39835	Arabidopsis thalia
873	39	23.5	268	21	AAAB8176	Thermus thermophil
874	39	23.5	270	21	AAU70148	Staphylococcus aur
875	39	23.5	282	14	AAAB3701	isoform of human o
876	39	23.5	288	22	AAAG82012	S. epidermidis ope
877	39	23.5	294	11	AAAB06435	Haloacetate dehalo
878	39	23.5	294	11	AAAB06435	Human olfactory an
879	39	23.5	295	14	ABP40798	Staphylococcus epi
880	39	23.5	296	14	AAAB30700	Human uterotrocin pr
881	39	23.5	299	22	AAU24536	Human olfactory re
882	39	23.5	299	22	AAAG71950	Human olfactory re
883	39	23.5	299	23	AAU95694	Human olfactory an
884	39	23.5	299	23	ABAB06643	G-protein-coupled
885	39	23.5	299	23	AAU85160	G-coupled olfactor
886	39	23.5	300	11	AAAB05161	Sequence of human
887	39	23.5	300	20	AAU05402	Human osteopontin.
888	39	23.5	300	21	AAAB19769	Human osteopontin.
889	39	23.5	300	22	AAAB30574	A human Eta-1/oste
890	39	23.5	300	23	AAU76591	Human osteopontin
891	39	23.5	303	11	AAAB05119	Intracellular alka
892	39	23.5	305	21	AAAB56627	Human prostate can
893	39	23.5	305	21	AAAG05315	Arabidopsis thalia
894	39	23.5	305	21	AAAG39834	Arabidopsis thalia
895	39	23.5	305	21	AAAG39834	Arabidopsis thalia
896	39	23.5	305	21	AAAG39834	Arabidopsis thalia
897	39	23.5	314	20	AAAB99779	Human osteopontin.
898	39	23.5	314	21	AAAB19770	Human osteopontin.
899	39	23.5	314	21	AAAB19770	Human osteopontin.
900	39	23.5	314	22	AAAB12663	Human osteopontin.
901	39	23.5	314	22	AAAB30573	A human Eta-1/oste
902	39	23.5	314	23	AAAB76043	Human osteopontin
903	39	23.5	317	21	AAAB40897	Human ORFX ORF661
904	39	23.5	317	23	ABP07634	Human ORFX protein
905	39	23.5	318	20	AAU08568	B. subtilis hydrol
906	39	23.5	322	23	AAAB21152	Human 52991 prodom
907	39	23.5	328	20	AAAB14581	Group B streptococ
908	39	23.5	328	22	ABAB4815	Drosophila melanog
909	39	23.5	329	22	AAU41574	Proprionbacterium
910	39	23.5	331	23	ABP28882	Streptococcus poly
911	39	23.5	337	22	AAU07601	Human tumour necro
912	39	23.5	338	22	ABG12191	Novel human diagno
913	39	23.5	353	21	AAU58152	penicillium funicu
914	39	23.5	353	22	AAAG10357	Novel human diagno
915	39	23.5	356	23	ABAB29386	Hereticidally activ
916	39	23.5	359	22	ABAB67964	Drosophila melanog
917	39	23.5	365	21	AAAB10674	A. thaliana PKBP-1
918	39	23.5	365	21	AAAG17165	Arabidopsis thalia
919	39	23.5	365	21	AAAG39583	Arabidopsis thalia
920	39	23.5	366	22	ABG29586	Novel human diagno
921	39	23.5	370	20	AAU06623	Human tyrosylprote
922	39	23.5	370	20	AAU06624	Mouse tyrosylprote
923	39	23.5	370	21	AAU84304	A human tyrosylpro
924	39	23.5	370	21	AAU84305	A murine tyrosylpr
925	39	23.5	380	22	AAAB69420	Amino acid sequenc
926	39	23.5	380	22	AAAB0385	Human gene 18 enco
927	39	23.5	380	23	ABG64570	Human albumin fusi
928	39	23.5	383	15	AAAB72790	Patatin coded by t
929	39	23.5	383	19	AAU74512	Amino acid sequenc
930	39	23.5	390	23	ABAB91492	Hereticidally activ
931	39	23.5	393	21	AAAG39833	Arabidopsis thalia
932	39	23.5	393	22	AAAB88461	Human membrane of
933	39	23.5	394	21	AAAG05314	Arabidopsis thalia
934	39	23.5	394	21	AAAG39979	Hereticidally activ
935	39	23.5	395	23	ABAB39290	Hereticidally activ
936	39	23.5	402	21	AAAG39978	Arabidopsis thalia
937	39	23.5	403	21	AAAG05313	Arabidopsis thalia
938	39	23.5	411	21	AAU54878	Proprionbacterium
939	39	23.5	417	23	ABP26857	Streptococcus poly
940	39	23.5	422	22	ABG07546	Novel human diagno
941	39	23.5	422	22	ABG21286	Novel human diagno
942	39	23.5	429	20	AAU37442	Chlamydia trachoma
943	39	23.5	434	22	AAAB79565	Corynebacterium q1
944	39	23.5	444	19	AAAB98306	H. pylori GHP 525
945	39	23.5	446	21	AAAG53797	Arabidopsis thalia
946	39	23.5	454	23	ABAB8331	Listeria monocytog
947	39	23.5	458	21	AAAB19869	Chimeric PCB decon
948	39	23.5	458	21	AAAB19869	Chimeric PCB decon
949	39	23.5	471	23	ABAB49151	Listeria monocytog
950	39	23.5	477	22	AAAG1294	C glutamicum prote
951	39	23.5	477	22	AAAB79802	Corynebacterium q1
952	39	23.5	477	23	ABAB2112	Hereticidally activ
953	39	23.5	500	23	AAAB21709	Human PKIN-4 prote
954	39	23.5	503	21	AAAB19126	Polypeptide isolat
955	39	23.5	513	23	AAAB19989	Protein of a human
956	39	23.5	516	23	ABAB1981	Hereticidally activ
957	39	23.5	520	22	AAAB94520	Human protein sequ
958	39	23.5	522	21	AAAG53796	Arabidopsis thalia
959	39	23.5	527	23	AAU98516	Novel extracellular

960	39	23.5	528	23	AAU72355	Neisseria meningit
961	39	23.5	536	22	ABB62486	Drosophila melanog
962	39	23.5	544	22	AAE09716	Human extracellular
963	39	23.5	544	23	ABP27253	Streptococcus poly
964	39	23.5	544	23	AAU72718	Human kinase h1418
965	39	23.5	547	23	AAE14507	Human catboxyleste
966	39	23.5	557	22	AAU78602	Human protein SFO
967	39	23.5	558	22	AAU03526	Human protein kina
968	39	23.5	568	22	ABG21986	Novel human diago
969	39	23.5	571	21	AB18914	A novel polypeptid
970	39	23.5	571	21	AAU71107	Human Hydrolyase pr
971	39	23.5	571	22	AAU12442	Human PRO1887 poly
972	39	23.5	571	22	AAE04101	Human gene 2 encod
973	39	23.5	571	23	ABG64341	Human albumin fusi
974	39	23.5	571	23	ABG84949	Human PRO1887 prot
975	39	23.5	572	23	AAU83696	Human PRO1887 prot
976	39	23.5	572	23	ABH77576	Physcomitrella pat
977	39	23.5	572	23	ABG55555	Human angiotensin
978	39	23.5	574	22	ABG58877	Drosophila melanog
979	39	23.5	580	19	AAU11093	H. pylori ORF hp6p
980	39	23.5	581	22	AAU33503	Novel human secret
981	39	23.5	583	19	AAU71506	Helicobacter polyp
982	39	23.5	593	22	AAU46350	H. pylori HPN191 p
983	39	23.5	605	21	AAU26942	Soybean auxin tran
984	39	23.5	606	22	ABG52667	Escherichia coli p
985	39	23.5	610	21	AAU53795	Arabidopsis thalia
986	39	23.5	612	22	AAU79786	Human protein SEQ
987	39	23.5	621	22	AAU94433	Human protein sequ
988	39	23.5	708	22	AAU48037	Proionibacterium
989	39	23.5	711	19	AAU55103	Streptococcus pneu
990	39	23.5	711	23	ABP54597	S. pneumoniae SP05
991	39	23.5	717	22	ABG14453	Drosophila melanog
992	39	23.5	721	22	ABG15463	Novel human diago
993	39	23.5	727	20	AAU27127	Human brain tissue
994	39	23.5	777	21	AAU99427	Human PRO1491 (ONQ
995	39	23.5	777	22	AAU29197	Human PRO polypept
996	39	23.5	777	22	AAU66176	Protein of the inv
997	39	23.5	783	19	AAU64560	Candida albicans o
998	39	23.5	783	21	AAU96164	Candida albicans o
999	39	23.5	800	22	ABG16023	Novel human diago
1000	39	23.5	807	22	AAU62729	Amino acid sequenc

ALIGNMENTS

RESULT 1
ID ABB81560 standard; Protein; 171 AA.

XX AC ABB81560;
XX DT 05-SEP-2002 (first entry)
XX DE Human high endothelial cell GLCNAC6ST portion SEQ ID NO:8.
XX KW Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GLCNAC6ST;
XX KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
XX KW ophthalmological.
XX OS Homo sapiens.
XX PN US2002061562-A1.
XX PD 23-MAY-2002.
XX PF 09-AUG-2001; 2001US-0927602.
XX PR 11-AUG-2000; 2000US-325773P.
XX PA (FUKU/) FUKUDA M N.
XX PA (AKAM/) AKAMA T O.
XX

PI Fukuda MN, Akama TO;
XX WPI: 2002-507643/54.
DR
XX
PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy
XX
PS Example 2; Fig 3C; 69pp; English.
XX
CC The present invention describes human corneal
CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
CC and has ophthalmological activity. (I) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (I) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents a portion of the human high endothelial cell
CC N-acetylglucosamine-6-sulfotransferase (GLCNAC6ST), which is given in
CC the exemplification of the present invention.
XX
SQ Sequence 171 AA:

Query Match 100.0%; Score 166; DB 23; Length 171;
Best Local Similarity 100.0%; Pred. No. 7.6e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSSGFVQGLFGQHPDVFYLMPEAWHY 29
DB 19 SWRSSGFVQGLFGQHPDVFYLMPEAWHY 47

RESULT 2
ID AAU11274 standard; Protein; 380 AA.

XX AC AAU11274;
XX DT 12-MAR-2002 (first entry)
XX DE Human L-selectin sulfotransferase-2 (LSS2-2) protein.
XX KW Human; beta1,3GNT; beta1,3-N-acetylglucosaminyltransferase; MECA-79;
XX KW L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;
XX KW ulcerative colitis; inflammatory skin disorder; psoriasis; lichen planus;
XX KW allergic contact dermatitis; lymphoma; chronic pneumonia; LSS2-2;
XX KW delayed-type hypersensitivity reaction; hyperplastic lymphus; antileuker;
XX KW antiinflammatory; antipsoriatic; antidiabetic; dermatological;
XX KW antiallergic.
XX OS Homo sapiens.
XX PN WO200185177-A1.
XX PD 15-NOV-2001.
XX PF 10-MAY-2001; 2001WO-US15452.
XX PR 11-MAY-2000; 2000US-0569320.
XX PA (BORN-) BORNHAM INST.
XX PA Fukuda M, Yeh J, Hiraoka N;
XX WPI: 2002-075226/10.
XX DR N-PSDB; AAS16947.
XX
PT New enzyme, useful for modifying acceptor molecule, comprises an
PT isolated L-selectin sulfotransferase-2 that directs expression of
PT L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or
PT intestinal GLCNAC 6-sulfotransferase
XX

PS Claim 21; Fig 4; 98pp; English.

XX The present invention provides a method of modifying an acceptor molecule

CC by contacting the acceptor with an isolated

CC beta1,3-N-acetylglucosaminyltransferase (beta1,3gnt) or an active

CC fragment, where beta1,3gnt directs expression of a MECA-79 antigen. The

CC invention also provides a method of treating or preventing an

CC L-selectin-mediated condition by reducing the expression or activity of a

CC beta1,3gnt that directs expression of a MECA-79 antigen. This can be done

CC by administering to the subject an oligosaccharide L-selectin antagonist

CC that inhibits binding of L-selectin to a MECA-79 antigen, for example by

CC administering antibody material that specifically binds beta1,3gnt,

CC and/or a beta1,3gnt antisense nucleic acid molecule. L-selectin

CC sulfotransferase-2 (LST-2) also directs MECA-79 antigen expression.

CC Alternatively, the expression or activity of LST-2 or its active

CC fragment can be reduced in combination with reducing the expression or

CC activity of beta1,3gnt. The method is useful for treating L-selectin

CC mediated conditions such as Crohn's disease and ulcerative colitis,

CC inflammatory disorders of the skin such as allergic contact dermatitis,

CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type

CC hypersensitivity reactions, diabetes and hyperplastic thymus. This

CC sequence represents human LST-2.

XX

Sequence 380 AA;

Query Match 100.0%; Score 166; DB 23; Length 380;

Best Local Similarity 100.0%; Pred. No. 1,9e-17;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SWRSGSFFVGQLFGQHPDVFYLMPEAMHV 29

|||||

DB 44 SWRSGSFFVGQLFGQHPDVFYLMPEAMHV 72

|||||

RESULT 3

AAV39918

ID AAV39918 standard; Protein: 386 AA.

XX

AC AAV39918;

XX

DT 08-DEC-1999 (first entry)

XX

DE Human glycosyl sulfotransferase-3 protein sequence.

XX

KW Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;

KW selectin binding interaction; inflammation; lymphocyte homing; human;

KW secondary lymph organ.

XX

OS Homo sapiens.

XX

PN WO9949018-A1.

XX

PD 30-SEP-1999.

XX

PF 26-FEB-1999; 99WO-US04316.

XX

PR 20-MAR-1998; 98US-0045284.

XX

PR 12-NOV-1998; 98US-0190911.

XX

PA (REGC) UNIV CALIFORNIA.

PA (SYNT) SYNTAX USA INC.

XX

PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;

XX

DR WPT; 1999-580442/49.

XX

DR N-PSDB; AA220792.

XX

PT Human and murine glycosyl sulfotransferase 3 and related

PT polynucleotides

XX

PS Claim 2; Fig 1; 59pp; English.

XX

XX This sequence is the human glycosyl sulfotransferase-3 (GST-3) of

CC the invention. The nucleic acid sequences, probes and primers derived

CC from these, proteins and antibodies are useful in detecting homologues.

CC The sequences, antibodies and methods are useful in the diagnosis and

CC treatment of diseases associated with selectin binding interactions,

CC including conditions associated with or resulting from the homing of

CC leukocytes to sites of inflammation and the normal homing of lymphocytes

CC to secondary lymph organs.

XX

Sequence 386 AA;

Query Match 100.0%; Score 166; DB 20; Length 386;

Best Local Similarity 100.0%; Pred. No. 2e-17;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SWRSGSFFVGQLFGQHPDVFYLMPEAMHV 29

|||||

DB 50 SWRSGSFFVGQLFGQHPDVFYLMPEAMHV 78

|||||

RESULT 4

AAV79219

ID AAV79219 standard; Protein: 386 AA.

XX

AC AAV79219;

XX

DT 19-JUN-2000 (first entry)

XX

DE Human transferase TRNSFS-11.

XX

XX

KW Transferase; TRNSFS-11; human; antitumor; cell proliferation;

KW gastrointestinal disorder; developmental disorder;

KW genetic disorder; neurological disorder; reproductive disorder;

KW smooth muscle disorder; immunological disorder; inflammation;

KW diagnosis; therapy; N-acetylglucosamine 6-O-sulfotransferase.

XX

OS Homo sapiens.

XX

PH Key

XX

FT Modified-site 121 Location/Qualifiers

FT Modified-site 107 /note= "potential O-phosphorylation"

FT Modified-site 217 /note= "potential O-phosphorylation"

FT Modified-site 252 /note= "potential O-phosphorylation"

FT Modified-site 364 /note= "potential O-phosphorylation"

FT Modified-site 380 /note= "potential O-phosphorylation"

FT Modified-site 35 /note= "potential O-phosphorylation"

FT Modified-site 50 /note= "potential O-phosphorylation"

FT Modified-site 81 /note= "potential O-phosphorylation"

FT Modified-site 287 /note= "potential O-phosphorylation"

FT Modified-site 243 /note= "potential O-phosphorylation"

FT Modified-site 30 /note= "potential O-phosphorylation"

FT Modified-site 308 /note= "potential N-glycosylation"

FT Modified-site 329 /note= "potential N-glycosylation"

FT Modified-site 7..23 /note= "potential N-glycosylation"

FT Domain

XX

PN WO200014251-A2.

XX

PD 16-MAR-2000.

PF 09-SEP-1999; 99WO-US20989.
XX
PR 10-SEP-1998; 98US-0150657.
PR 04-NOV-1998; 98US-0186779.
PR 11-MAY-1999; 99US-0133642.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
PI Hillman JL, Azimzai Y;
XX
DR WPI, 2000-25696/22.
DR N-PSDB; AA294211.
XX
PT Human transferase proteins useful for preventing, diagnosing and
PT treating cancers and developmental, gastrointestinal, genetic,
PT immunological, neurological, reproductive and smooth muscle disorders -
XX
PS Claim 1; Page 90-91; 113pp; English.
XX
CC The present sequence is that of human transferase TRNSFS-11, 1 of
CC 15 claimed novel human transferase proteins of the invention (see
CC AAY79209-23). The sequence was deduced from a cDNA clone (see
CC AA294211) isolated from a gallbladder library. It shows homology to
CC mouse N-acetylglucosamine 6-O-sulfotransferase. TRNSFS-11 is
CC expressed in dermatologic and gastrointestinal tissues,
CC especially those associated with inflammation and cell
CC proliferation. The new human transferases and polynucleotides can
CC be used in the diagnosis, prevention and treatment of cancer,
CC developmental disorders, gastrointestinal disorders, genetic
CC disorders, immunological disorders, neurological disorders,
CC reproductive disorders, and smooth muscle disorders. The
CC polypeptides can also be used to raise antibodies, and to screen
CC for agonists and antagonists of transferase activity.
XX
SQ Sequence 386 AA;

Query Match 100.0%; Score 166; DB 21; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVGLFGQHPDYFLMEPAMHV 29
Db 50 SMRSGSFVGLFGQHPDYFLMEPAMHV 78
|||||
RESULT 5
AA93309 standard; Protein: 386 AA.
XX
AC AAM93309;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 2817.
XX
KM Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELT-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI, 2001-52425/58.
DR N-PSDB; AA94229.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 2817; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' and 3' ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesized by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 386 AA;

Query Match 100.0%; Score 166; DB 22; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVGLFGQHPDYFLMEPAMHV 29
Db 50 SMRSGSFVGLFGQHPDYFLMEPAMHV 78
|||||
RESULT 6
AA93919 standard; Protein: 388 AA.
XX
AC AAY39919;
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 protein sequence.
XX
KM Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KM selectin binding interaction; inflammation; lymphocyte homing; mouse;
KM secondary lymph organ.
XX
OS Mus sp.
XX
PN W09949018-A1.
XX
PD 30-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIV CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistrup A, Rosen SD, Tangemann K, Hammerich S;
XX
DR WPI, 1999-58044/2/49.
DR N-PSDB; AA20793.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
PT polynucleotides -
XX
PS Claim 2; Fig 4; 59pp; English.
XX
CC This sequence is the mouse glycosyl sulfotransferase-3 (GST-3) of

CC the invention. The nucleic acid sequences, probes and primers derived
 CC from these, proteins and antibodies are useful in detecting homologues.
 CC The sequences, antibodies and methods are useful in the diagnosis and
 CC treatment of diseases associated with selectin binding interactions,
 CC including conditions associated with or resulting from the homing of
 CC leukocytes to sites of inflammation and the normal homing of lymphocytes
 CC to secondary lymph organs.
 CC
 CC

XX Sequence 388 AA;

Query Match

Best Local Similarity 100.0%; Score 166; DB 20; Length 388;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 29
 DB 49 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 77

RESULT 7

ABB81558

ID ABB81558 standard; Protein: 169 AA.

AC ABB81558;

XX

DT 05-SEP-2002 (first entry)

XX

DE Human corneal N-acetylglucosamine-6-sulfotransferase portion SEQ:6.

XX

KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme; G1CNA6ST;

KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;

KW ophthalmological.

XX

OS Homo sapiens.

XX

PN US2002061562-A1.

XX

PD 23-MAY-2002.

XX

PF 09-AUG-2001; 2001US-0927602.

XX

PR 11-AUG-2000; 2000US-325773P.

XX

PA (FUKU/) FUKUDA M N.

XX

FA (AKAM/) AKAMA T O.

XX

PI Fukuda MN, Akama TO;

XX

WPI; 2002-507643/54.

PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX
 XX Example 2; Fig 3C; 69pp; English.

CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human corneal N-acetylglucosamine-
 CC 6-sulfotransferase (G1CNA6ST), which is given in the exemplification
 CC of the present invention.
 CC
 CC

XX Sequence 169 AA;

Query Match

Best Local Similarity 96.4%; Score 160; DB 23; Length 169;
 Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 29
 DB 16 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 44

RESULT 8
 AAY72638

ID AAY72638 standard; Protein: 395 AA.

AC AAY72638;

XX

DT 02-MAY-2001 (first entry)

XX

DE Mouse glycosyl sulfotransferase-4 (GST-4).

XX

KW Mouse; glycosyl sulfotransferase-4; GST-4; immunosuppressive;

KW therapy; selectin binding inhibitor; gene therapy; inflammation;

KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;

KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;

KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;

KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;

KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;

KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;

KW asthma; hypersensitivity; rheumatic fever; tissue rejection;

KW chromosome 8E1.

XX

OS Mus musculus.

XX

PN W0200106015-A1.

XX

PD 25-JAN-2001.

XX

PF 19-JUL-2000; 2000MO-US19741.

XX

PR 20-JUL-1999; 99US-0144694.

XX

PR 13-JUL-2000; 2000US-0593828.

XX

PA (REGC) UNIV CALIFORNIA.

XX

PI Rosen SD, Lee JK, Hammerich S;

XX

PT New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
 PT diagnostic and therapeutic agent screening applications -
 XX
 XX Claim 3; Fig 2; 128pp; English.

CC The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4
 CC gene is found on chromosome 8E1.
 CC GST is a type 2 membrane protein useful for inhibiting a binding event
 CC between a selectin and a selectin ligand, which comprises contacting the
 CC selectin with a non-sulphated selectin ligand, GST and a small molecular
 CC agent that inhibits the sulphation mediated binding event. GST is useful in gene
 CC therapy to treat disorders such as acute or chronic inflammation,
 CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
 CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
 CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
 CC disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
 CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
 CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
 CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
 CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
 CC during transplantation.
 CC
 CC

XX Sequence 395 AA;

Query Match

Best Local Similarity 96.4%; Score 160; DB 22; Length 395;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 29
 |||
 DB 48 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 76

RESULT 9

ID AAY72640
 AAY72640 standard; Protein: 395 AA.

AC AAY72640;

DT 02-MAY-2001 (first entry)

DE Human glycosyl sulfotransferase-4beta (GST-4beta).

XX Human; glycosyl sulfotransferase-4beta; GST-4beta; immunosuppressive;
 XX therapy; selectin binding inhibitor; gene therapy; inflammation;
 KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
 KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
 KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
 KW chromosome 16q23.1.

OS Homo sapiens.

PN MO200106015-A1.

PD 25-JAN-2001.

PF 19-JUL-2000; 2000MO-US19741.

PR 20-JUL-1999; 99QS-0144694.

PR 13-JUL-2000; 2000US-0593826.

XX (REGC) UNIV CALIFORNIA.

PA Rosen SD, Lee JK, Hemmerich S;

PI WPI: 2001-138471/14.

DR N-PSDB; AAD02697, AAD02700.

PT New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for

PS diagnostic and therapeutic agent screening applications

XX Claim 3; Fig 4B; 128bp; English.

CC The present sequence is human glycosyl sulfotransferase-4beta (GST-4
 CC beta). GST-4 gene is found on chromosome 16q23.1.
 CC GST is a type 2 membrane protein useful for inhibiting a binding event
 CC between a selectin and a selectin ligand, which comprises contacting the
 CC selectin with a non-sulphated selectin ligand, GST and a small molecular
 CC agent that inhibits the sulphation activity of GST. GST is also useful
 CC in inhibiting a selectin mediated binding event. GST is useful in gene
 CC therapy to treat disorders such as acute or chronic inflammation,
 CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
 CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
 CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
 CC disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
 CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
 CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
 CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
 CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
 CC during transplantation.

XX Sequence 395 AA;

Query Match 96.4%; Score 160; DB 22; Length 395;
 Best Local Similarity 96.6%; Pred. No. 1.8e-16;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 29
 |||
 DB 48 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 76

RESULT 10

ID ABB81554
 ABB81554 standard; Protein: 395 AA.

AC ABB81554;

DT 05-SEP-2002 (first entry)

DE Human corneal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:2.

XX Human; N-acetylglucosamine-6-sulfotransferase; enzyme: GlcNAc6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological; chromosome 16q22.

OS Homo sapiens.

PN US2002061562-A1.

PD 23-MAY-2002.

PF 09-AUG-2001; 2001US-0927602.

PR 11-AUG-2000; 2000US-325773P.

PA (FUKU/) FUKUDA M N.

PA (AKAM/) AKAMA T O.

PI Fukuuda MN, Akama TO;

DR WPI: 2002-507643/54.

DR N-PSDB; ABN89506.

PT New nucleic acid encoding corneal

PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,

PT monitoring and diagnosis of macular corneal dystrophy

PS Claim 13; Fig 1A-D; 69bp; English.

CC The present sequence represents human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratinoplasty or keratectomy.

XX Sequence 395 AA;

Query Match 96.4%; Score 160; DB 23; Length 395;
 Best Local Similarity 96.6%; Pred. No. 1.8e-16;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 29
 |||
 DB 48 SMRSGSFVGOLEFGQHPDVFYLMPEAMHV 76

RESULT 11

ID ABB81555
 ABB81555 standard; Protein: 395 AA.

AC ABB81555;

DT 05-SEP-2002 (first entry)

DE Consensus N-acetylglucosamine-6-sulfotransferase SEQ ID NO:3.

XX Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GLCNAC6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Synthetic.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 10
 FT /label= Ala, Thr, Val
 FT Misc-difference 13
 FT /label= Ala, Val, Ser
 FT Misc-difference 20
 FT /label= Phe, Cys, Gly
 FT Misc-difference 39
 FT /label= Ala, Asp, Glu
 FT Misc-difference 96
 FT /label= Val, Met, Ile
 FT Misc-difference 142
 FT /label= Ala, Thr, Asn
 FT Misc-difference 147
 FT /label= Ala, Asp, Glu
 FT Misc-difference 159
 FT /label= Thr, Ser, Gly
 FT Misc-difference 238
 FT /label= Gly, His, Arg
 FT Misc-difference 294
 FT /label= Ser, Thr, Gly
 FT Misc-difference 371
 FT /label= Ala, Thr, Ser
 FT Misc-difference 380
 FT /label= Leu, Pro, Met
 FT Misc-difference 382
 FT /label= Gly, His, Ser
 FT Misc-difference 384
 FT /label= Thr, Ser, Lys
 FT Misc-difference 390
 FT /label= Ala, Glu
 FT Misc-difference 391
 FT /label= Ser, Lys
 FT Misc-difference 392
 FT /label= His, Gln
 FT Misc-difference 394
 FT /label= Arg, Glu
 FT Misc-difference 395
 FT /label= Asn, Ser
 FT
 XX US2002061562-A1.
 XX
 XX 23-MAY-2002.
 XX
 XX 09-AUG-2001; 2001US-0927602.
 XX
 XX 11-AUG-2000; 2000US-325773P.
 XX
 XX (FUKU/) FUKUDA M N.
 XX (AKAM/) AKAMA T O.
 XX
 XX Fukuda MN, Akama TO;
 XX
 XX WPI; 2002-507643/54.
 XX
 XX New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX Example 5; Fig 2A-B; 69pp; English.
 XX
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for

CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a consensus N-acetylglucosamine-6-sulfotransferase
 CC which is given in the exemplification of the present invention.
 CC
 XX
 SO Sequence 395 AA:
 QY 1 SWRSGSSFVGOIRFGQHPDYFLMEPAWV 29
 Db 48 SWRSGSSFVGOIRFGQHPDYFLMEPAWV 76
 |||||
 RESULT 12
 ID AAE15438
 AC AAE15438; Protein; 395 AA.
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Human drug metabolizing enzyme (DME)-5.
 XX
 KW Human; drug metabolizing enzyme; gene therapy; autoimmune disorder;
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;
 KW cancer; endocrine disorder; hypochalimus disorder; pituitary disorder;
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;
 KW DME-5.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..32
 FT /label= Signal-peptide
 FT Protein 33..395
 FT /note= "Human mature DME-5 protein"
 FT Peptide 1..35
 FT /label= Signal-peptide
 FT Protein 36..395
 FT /note= "Human mature DME-5 protein"
 FT
 XX WO200179468-A2.
 XX
 XX 25-OCT-2001.
 XX
 XX 12-APR-2001; 2001WO-US11869.
 XX
 XX 13-APR-2000; 2000US-197590P.
 XX 19-APR-2000; 2000US-198403P.
 XX 28-APR-2000; 2000US-200185P.
 XX 05-MAY-2000; 2000US-202234P.
 XX 11-MAY-2000; 2000US-203509P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Pollicky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM;
 PI Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS;
 PI Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 PI Au-Young J;
 XX
 XX WPI; 2002-066363/09.
 XX N-PDB; AAD4670.
 XX


```

PA (AKAM//) AKAMA T O.
XX
PI Fukuda MN, Akama TO;
XX
DR WPI; 2002-507643/54.
XX
PT New nucleic acid encoding corneal
PT N-acetylglicosamine-6-sulfotransferase, useful for treatment.
PT monitoring and diagnosis of macular corneal dystrophy -
XX
PS Example 5; Page 24-25; 69pp; English.
XX
CC The present invention describes human corneal
CC N-acetylglicosamine-6-sulfotransferase (I), which is able to catalyze
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
CC and has ophthalmological activity. (I) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (I) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents mouse interstitial N-acetylglicosamine-6-
CC sulfotransferase, which is given in comparison with (I) in the
CC exemplification of the present invention.
XX
SQ Sequence 418 AA:
XX
Query Match 96.4%; Score 160; DB 23; Length 418;
Best Local Similarity 96.6%; Pred. No. 1,9e-16;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SWRSGSFFVGLFGQHPDYFLMEPAMHY 29
|||||
DB 72 SWRSGSFFVGLFGQHPDYFLMEPAMHY 100
|||||

RESULT 15
ABB81559
ID ABB81559 standard; Protein; 169 AA.
XX
AC ABB81559;
XX
DT 05-SEP-2002 (first entry)
DE Human intestinal N-acetylglicosamine-6-sulfotransferase portion SEQ:7.
XX
KW Human; N-acetylglicosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological.
XX
OS Homo sapiens.
XX
EN US2002061562-A1.
XX
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU//) FUKUDA M N.
PA (AKAM//) AKAMA T O.
PI Fukuda MN, Akama TO;
XX
DR WPI; 2002-507643/54.
XX
PT New nucleic acid encoding corneal
PT N-acetylglicosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy -
XX
PS Example 2; Fig 3C; 69pp; English.
XX
CC The present invention describes human corneal

```

```

CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
CC and has ophthalmological activity. (I) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (I) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents a portion of the human intestinal
CC N-acetylglucosamine-6-sulfotransferase (GlcNAc6ST), which is given in
CC the exemplification of the present invention.
XX
XX
XX Sequence 169 AA;
XX
XX
XX Query Match 94.6%; Score 157; DB 23; Length 169;
XX Best Local Similarity 93.1%; Pred. No. 1.9e-16;
XX Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 SWRSGSFFVGOLFQHPDVFYLMEDAMHY 29
XX |||||:||||:||||:||||:||||:
XX 16 SWRSGSFFLGOLFSGHPDVFYLMEDAMHY 44
XX
XX
XX RESULT 16
XX AA772639
XX ID AA772639 standard; Protein: 390 AA.
XX AC
XX AA772639;
XX
XX 02-MAY-2001 (first entry)
XX
XX Human glycosyl sulfotransferase-4alpha (GST-4alpha) .
XX
XX Human; glycosyl sulfotransferase-4alpha; GST-4alpha; immunosuppressive;
XX therapy; selectin binding inhibitor; gene therapy; inflammation;
XX systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
XX polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
XX glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenailitis;
XX Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
XX demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
XX myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
XX asthma; hypersensitivity; rheumatic fever; tissue rejection;
XX chromosome 16q23.1.
XX
XX Homo sapiens.
XX
XX WO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-051974.1.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerlich S;
XX
XX WPI; 2001-138471/14.
XX N-PSDB; AAD02697, AAD02698, AAD02699.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications -
XX
XX Claim 3; Fig 1; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
XX alpha). GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene

```

CC therapy to treat disorders such as acute or chronic inflammation,
CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
CC disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
CC during transplantation.
CC Note: The present sequence is also shown in sequence listing (page
CC no: 56) but lacks four nucleotides at its 3' end.
XX
SQ Sequence 390 AA;
XX
Query Match 94.6%; Score 157; DB 22; Length 390;
Best Local Similarity 93.1%; Pred. No. 5.1e-16;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMRSGSSFVGQFGQHPDVFYIMEPAHV 29
DB 49 SMRSGSSFLGQFSGHPDVFYIMEPAHV 77
RESULT 17
ABB81556
ID ABB81556 standard; Protein; 390 AA.
XX
AC ABB81556;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human intestinal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:4.
XX
KM Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
KM corneal; sulfatase; keratan sulfate; macular corneal dystrophy; MCD;
KM ophthalmological.
XX
OS Homo sapiens.
XX
PN US2002061562-A1.
XX
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU/) FUKUDA M. N.
PA (AKAM/) AKAMA T. O.
XX
PI Fukuda MN, Akama TO;
XX
PI WPI; 2002-507643/54.
XX
DR
XX
PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy -
XX
PS Example 5; Fig 2A-B; 69pp; English.
XX
CC The present invention describes human corneal
CC N-acetylglucosamine-6-sulfotransferase (1), which is able to catalyse
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (1) is located to chromosome 16q22,
CC and has ophthalmological activity. (1) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (1) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents human intestinal N-acetylglucosamine-6-
CC sulfotransferase, which is given in comparison with (1) in the
CC exemplification of the present invention.

SQ Sequence 390 AA;
XX
Query Match 94.6%; Score 157; DB 23; Length 390;
Best Local Similarity 93.1%; Pred. No. 5.1e-16;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMRSGSSFVGQFGQHPDVFYIMEPAHV 29
DB 49 SMRSGSSFLGQFSGHPDVFYIMEPAHV 77
RESULT 18
ABB41947
ID ABB41947 standard; Protein; 418 AA.
XX
AC ABB41947;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1711 polypeptide sequence SEQ ID NO:3422.
XX
KM Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KM vulnery; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
KM anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM cholesterol ester storage; systemic lupus erythematosus; infection;
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM allergy; aplastic anaemia; nocturnal hemoglobinuria; burn; wound;
KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
KM thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CORA-) CORAGEN CORP.
XX
PI Shinketsu RA, Leach M;
XX
PI WPI; 2000-602362/57.
XX
DR N-PSDB; AAC76156.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 2599-2600; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in ABB40237 to ABB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritis; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antineumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The


```

RESULT 21
XX AAY31657 standard; Protein; 484 AA.
ID AAY31657
XX
AC AAY31657;
XX
DT 09-NOV-1999 (first entry)
XX
DE Human N-acetylglucosamine-6-O-sulfotransferase.
XX
KW N-acetylglucosamine-6-O-sulfotransferase; human; GLYCAM-1;
XX L-selectin ligand.
XX
OS Homo sapiens.
XX
PN EP943688-A2.
XX
PD 22-SEP-1999.
XX
PF 04-MAR-1999; 99EP-0301530.
XX
PR 24-JUN-1998; 98JP-0177844.
XX
PR 05-MAR-1998; 98JP-0054007.
XX
PA (SEK) SEIKAGAKU CORP.
XX
PI Habuchi O, Kadomatsu K, Kanagaki R, Muramatsu H;
PI Muramatsu T, Uchimura K;
XX
DR WPI; 1999-520337/44.
DR N-PSDB; AAX87821.
XX
PT New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PT for synthesis of sugar chains, e.g. GLYCAM-1
XX
PS Claim 2; Page 28-30; 41pp; English.
XX
CC The present sequence represents human N-acetylglucosamine-6-O-
CC sulfotransferase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC  $\beta$ 1-3 =  $\beta$ 1-3 glycosidic linkage, and  $\beta$ 1-4 =  $\beta$ 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GLYCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87821) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GLYCAM-1 using transformants which harbour
CC the DNA.
XX
SQ Sequence 484 AA:
XX
Query Match 74.7%; Score 124; DB 20; Length 484;
Best Local Similarity 69.0%; Pred. No. 9.2e+11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SWRSSSFVGQIFGQHPDYFLTEPAMV 29
Db 126 TWRSSSFEGELFNQNPVEFLTEPAMV 154
XX
RESULT 22
XX AAB95367
ID AAB95367 standard; Protein; 530 AA.
XX
AC AAB95367;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17679.

```

```

XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300252.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 17679; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 530 AA:
XX
Query Match 74.7%; Score 124; DB 22; Length 530;
Best Local Similarity 69.0%; Pred. No. 1e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SWRSSSFVGQIFGQHPDYFLTEPAMV 29
Db 172 TWRSSSFEGELFNQNPVEFLTEPAMV 200
XX
RESULT 23
XX AA069414
ID AA069414 standard; Protein; 531 AA.
XX
AC AA069414;
XX
DT 30-JAN-2002 (first entry)
XX
DE

```


XX (SEK) SEIKAGAKU KOGYO CO LTD.
 XX
 XX Fukuoka M, Habuchi O;
 XX
 DR WPI: 1998-288750/26.
 DR N-PSDB: AAV36418.
 XX
 PT Keratan sulphate 6-sulpho-transferase polypeptide - transfers
 PT sulphate from sulphate donor to galactose 6-hydroxy group etc.
 XX
 PS Claim 4; Page 15-16; 21pp; English.
 CC This sequence produces the protein KSGal6ST (Keratan sulphate
 CC 6-sulphotransferase), which has a molecular weight of about 46700. The
 CC cDNA sequence of this protein was obtained by radiolabelling the cDNA of
 CC chick chondroitin 6-sulphotransferase (C6ST) and using this as a probe in
 CC a random oligonucleotide-primed labelling method. Human foetal brain
 CC cDNA was inserted into a phage lambda gtl cloning vector whereby the
 CC clones containing the KSGal6ST were obtained by hybridization using the
 CC prepared probe. The positive clones were subcloned into a recombinant
 CC expression vector and used to transform COS-7 cells, from which cells
 CC expressing KSGal6ST can be selected. The phage cDNA inserts were
 CC isolated and subcloned into a Bluescript plasmid. Deletion clones were
 CC then prepared from which both strands were sequenced by the Sanger
 CC method. The KSGal6ST of the invention transfers the sulphate from a
 CC sulphate donor to galactose 6-OH groups in keratan sulphate, but does not
 CC transfer sulphate to chondroitin, chondroitin sulphate A or C, dermatan
 CC sulphate or CDSNS heparin.
 CC
 SQ Sequence 411 AA;
 Query Match 69.9%; Score 116; DB 19; Length 411;
 Best Local Similarity 81.5%; Pred. No. 1.4e-09;
 Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 3 RSGSFFVQGLFGQHPDVFYLMPEAMHV 29
 DB 70 RSGSFFVQGLFGQHPDVFYLMPEAMHV 96
 RESULT 26
 ABB81563
 ID ABB81563 standard; Protein; 174 AA.
 XX
 XX ABB81563;
 XX
 DT 05-SEP-2002 (first entry)
 DE Human chondroitin-6-sulphotransferase portion SEQ ID NO:11.
 XX
 KW Human; N-acetylglucosamine-6-sulphotransferase; enzyme; G1NAC6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 XX
 FN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 09-AUG-2001; 2001US-0927602.
 XX
 PR 11-AUG-2000; 2000US-325773P.
 XX
 PA (FUKU/) FUKUDA M N.
 PA (AKAM/) AKAMA T O.
 XX
 PI Fukuoka MN, Akama TO;
 DR WPI: 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal

PT N-acetylglucosamine-6-sulphotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX
 XX Example 2; Fig 3C; 69pp; English.
 CC
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulphotransferase (I), which is able to catalyse
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human chondroitin-6-
 CC sulphotransferase, which is given in the exemplification of the present
 CC invention.
 CC
 SQ Sequence 174 AA;
 Query Match 60.8%; Score 101; DB 23; Length 174;
 Best Local Similarity 59.3%; Pred. No. 1.1e-07;
 Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 QY 3 RSGSFFVQGLFGQHPDVFYLMPEAMHV 29
 DB 22 RTGSSFFVGEFFNQGNIFYLFEPMLHI 48
 RESULT 27
 AAM06480
 ID AAM06480 standard; Protein; 458 AA.
 XX
 XX AAM06480;
 AC
 XX
 DT 25-FEB-1997 (first entry)
 DE Chick chondroitin 6-sulphotransferase.
 XX
 KW Chondroitin 6-sulphotransferase; C6ST; chondroitin sulphate;
 KW proteoglycan.
 XX
 OS Gallus sp.
 XX
 FH Key Location/Qualifiers
 FT 1..33
 FT Peptide /label= Sig_peptide
 FT 34..458
 FT Protein /label= Mat.protein
 FT /note= "preferred polypeptide (Claim 5)"
 FT 1..458
 FT Protein /note= "preferred polypeptide (Claim 3)"
 FT 20..458
 FT Protein /note= "preferred polypeptide (Claim 4)"
 FT 38..187
 FT Protein /note= "preferred polypeptide (Claim 6)"
 FT 24..37
 FT Domain /label= Transmembrane_domain
 FT 62..64
 FT Modified-site /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 73..75
 FT Modified-site /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 95..97
 FT Modified-site /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 236..238
 FT Modified-site /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 399..401
 FT Modified-site /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 443..445
 FT Modified-site

```

FT      /label= Glycosylation
XX      /note= "potential N-glycosylation site"
XX
XX      EP745668-A2.
XX
XX      04-DEC-1996.
XX
XX      30-MAY-1996; 96EP-0303887.
XX
XX      31-MAY-1995; 95JP-0134358.
XX
XX      (SEK ) SEIKAGAKU KOGYO CO LTD.
XX
XX      Fukuta M, Habuchi O;
XX
XX      WPI; 1997-013692/02.
XX
XX      N-PSDB; AAT45037.
XX
XX      DNA encoding chondroitin 6-sulphotransferase - for recombinant
XX      production of C6ST, for use in industrial processes
XX
XX      Claim 2; Page 18-20; 30pp; English.
XX
XX      Chick embryo chondrocyte chondroitin 6-sulphotransferase (C6ST)
XX      (AAW06480) is capable of catalysing the transfer of a sulphate group
XX      from 3'-phosphoadenosine 5' phosphosulphate to the hydroxyl group
XX      at the C-6 position of the N-acetylgalactosamine residue of
XX      chondroitin. Its amino acid sequence was deduced from a cDNA clone
XX      (AAI45037) isolated from a chick embryo chondrocyte cDNA library.
XX      Isolation of the cDNA allows the mass prodn. of C6ST in transformed
XX      host cells. Recombinant C6ST polypeptides are useful for analysing
XX      the activities of chondroitin sulphate and for modifying its
XX      function, and for raising antibodies.
XX
XX      Sequence 458 AA;
XX
XX      Query Match 60.8%; Score 101; DB 18; Length 458;
XX      Best Local Similarity 59.3%; Pred. No. 3.4e-07;
XX      Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
XX
XX      QY 3 RSGSFFVGQIFGHPDVFYIMEPAMHV 29
XX      122 RTGSSFFVGEEFNQGNIFYLEPLWHI 148
XX
XX      RESULT 28
XX      AAW52863
XX      ID AAW52863 standard; Protein: 479 AA.
XX
XX      AC AAW52863;
XX
XX      DT 30-JUN-1998 (first entry)
XX
XX      DE Glycosaminoglycan sulphotransferase protein.
XX
XX      KW Glycosaminoglycan sulphotransferase; C6ST; N-acetylglactosamine;
XX      galactose; glycosaminoglycan; chondroitin; keratan sulphate; diagnostic.
XX
XX      OS Homo sapiens.
XX
XX      PN EP821066-A1.
XX
XX      PD 28-JAN-1998.
XX
XX      PE 22-JUL-1997; 97EP-0305476.
XX
XX      PR 24-JUL-1996; 96JP-0195063.
XX
XX      PA (SEK ) SEIKAGAKU KOGYO CO LTD.
XX
XX      PI Fukuta M, Habuchi O;
XX
XX      WPI; 1998-102622/10.

```

```

XX      Human glycosaminoglycan sulpho:transferase protein - catalyses
XX      PT 6-sulphation of chondroitin to produce chondroitin sulphate for use
XX      in pharmaceutical(s)
XX
XX      Claim 3; Page 21-23; 27pp; English.
XX
XX      The human glycosaminoglycan sulphotransferase (C6ST) is a type II
XX      membrane protein derived from foetal brain tissue with a molecular weight
XX      of 50-55 kD. The protein transfers sulphate groups from a donor to the
XX      N-acetylgalactosamine residue or galactose residue of glycosaminoglycan.
XX      C6ST displays substrate specificity transferring the sulphate group to
XX      the hydroxyl group position at C-6 of the N-acetylgalactosamine residue
XX      of chondroitin and the hydroxyl group position at C-6 of the galactose
XX      residue of keratan sulphate. The protein is useful for studying the
XX      function of chondroitin sulphate and can provide chondroitin sulphate
XX      that may be useful in pharmaceuticals. Both the protein and DNA might
XX      be useful for treating or diagnosing diseases attributable to low C-6
XX      sulphation of chondroitin N-acetylgalactosamine residues.
XX
XX      Sequence 479 AA;
XX
XX      Query Match 60.8%; Score 101; DB 19; Length 479;
XX      Best Local Similarity 59.3%; Pred. No. 3.5e-07;
XX      Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
XX
XX      QY 3 RSGSFFVGQIFGHPDVFYIMEPAMHV 29
XX      142 RTGSSFFVGEEFNQGNIFYLEPLWHI 168
XX
XX      RESULT 29
XX      ABB64513
XX      ID ABB64513 standard; Protein: 315 AA.
XX
XX      AC ABB64513;
XX
XX      DT 26-MAR-2002 (first entry)
XX
XX      DE Drosophila melanogaster polypeptide SEQ ID NO 20331.
XX
XX      KW Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical.
XX
XX      OS Drosophila melanogaster.
XX
XX      PN W0200171042-A2.
XX
XX      PD 27-SEP-2001.
XX
XX      PE 23-MAR-2001; 2001WO-US09231.
XX
XX      PR 23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX
XX      PA (PEKE ) PE CORP NY.
XX
XX      PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX      WPI; 2001-656860/75.
XX
XX      N-PSDB; ABL08616.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
XX      Disclosure; SEQ ID NO 20331; 21pp + Sequence Listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX      CC useful in developmental biology and in elucidating cell signalling and
XX      CC cell-cell interactions in higher eukaryotes for the development of
XX      CC insecticides, therapeutics and pharmaceutical drugs. The invention

```

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS7072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 315 AA;

Query Match 42.8%; Score 71; DB 22; Length 315;
Best Local Similarity 52.0%; Pred. No. 0.01;
Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 SMRSGSFVQGLFGQHPDVYFIMEP 25
| | | | | : | : | | | |
Db 191 SMRSGSFVQGLFGQHPDVYFIMEP 215

RESULT 30
AAW08432
ID AAW08432 standard; Protein; 372 AA.
XX

AAW08432;

DT 25-FEB-1997 (first entry)

DE Human interleukin-1 converting enzyme homologue ICEY.

KM Interleukin-1 converting enzyme homologue; ICEY; cysteine protease;

KM TNP-1; inflammation; anti-inflammatory; apoptosis; diagnosis;

KM therapy.

OS Homo sapiens.

XX

PN WO9638569-A1.

XX

PD 05-DEC-1996.

XX

PF 30-MAY-1996; 96WO-US08150.

XX

PR 31-MAY-1995; 95US-0443865.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Braxton SM, Delegeane AM, Diep D;

XX

DR WPI, 1997-034381/03.

DR N-PSDB; AAT42675.

XX

PT DNA encoding interleukin-1 converting enzyme homologue (ICEY) -

PT useful for diagnosis and detection of related sequences and

PT inhibitors

XX

PS Claim 1; Page 31-32; 46pp; English.

XX

CC Human interleukin-1 converting enzyme homologue (ICEY) (AAW08432) is
CC a cysteine protease with pro-inflammatory activity. Excessive
CC expression may lead to tissue damage or destruction. Its amino
CC acid sequence was deduced from a DNA clone (AAT42675) derived from
CC phorbol ester- and endotoxin-activated monocytic leukaemic THP-1
CC cells. Recombinant ICEY can be produced in transformed host cells
CC and used in diagnostic tests for ICEY, to screen for inhibitors
CC useful as anti-inflammatory, and to raise diagnostic antibodies.
XX
SQ Sequence 372 AA;

RESULT 31
AAG92219
ID AAG92219 standard; Protein; 344 AA.
XX
AC AAG92219;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 5973.
XX
KM Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI, 2001-376931/40.
DR N-PSDB; AAG67438.
XX
PT Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 17; SEQ ID NO: 5973; 246pp + Sequence Listing; English.
XX
SQ The present invention provides a number of nucleotide and protein
SQ sequences from the Corynebacterium glutamicum. These
SQ sequences are useful for identifying the mutation point of a gene derived from a
SQ mutant of corynebacterium, measuring expression amount and
SQ analysing the expression profile or expression pattern of a gene derived
SQ from Corynebacterium, and identifying a homologue of a gene derived
SQ from Corynebacterium. Corynebacterium are useful for producing
SQ amino acids, nucleic acids, vitamins, saccharides and organic acids,
SQ particularly L-lysine. The present sequence is a protein described
SQ in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 344 AA;

Query Match 30.1%; Score 50; DB 22; Length 344;
Best Local Similarity 47.8%; Pred. No. 22;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

Qy 8 FVQGLFGQHPD-VFYLMEPANH 28
| : | | | | | | | | | |
Db 187 FVEELKDHDPAPVYLSTGANN 209

RESULT 32
AAE22917
ID AAE22917 standard; Protein; 791 AA.
AC AAE22917;
XX
XX

DT 09-AUG-2002 (first entry)

XX Human transporter and ion channel (TRICH) 16.

DE Human transporter and ion channel (TRICH) 16.

XX Human: transporter and ion channel; TRICH; transport disorder;

KM diabetes mellitus; angina; Alzheimer's disease; neurological disorder;

KM stroke; Huntington's disease; meningitis; muscle; myocarditis; cancer;

KM infectious myositis; arrhythmia; asthma; immunological; gene therapy;

KM acquired immunodeficiency syndrome; AIDS; allergy; atherosclerosis;

KM cell proliferative disorder; cerebroprotective; cirrhosis; hepatitis;

KM transgenic; neuroprotective; anticonvulsant; nootropic; cytostatic;

KM antiinflammatory; hepatotropic; psoriasis.

OS Homo sapiens.

XX Location/Qualifiers

XX Key 177..193

XX Domain /label=Transmembrane_domain

XX Domain 348..366

XX Domain /label=Transmembrane_domain

XX Modified-site 393..429

XX /note="ATPase phosphorylation site"

XX Modified-site 417..471

XX /note="ATPase phosphorylation site"

XX Modified-site 431..449

XX /note="ATPase phosphorylation site"

XX Modified-site 575..585

XX /note="ATPase phosphorylation site"

XX Modified-site 644..684

XX /note="ATPase phosphorylation site"

XX WO200222684-A2.

XX 21-MAR-2002.

XX 14-SEP-2001; 2001WO-US28938.

XX 15-SEP-2000; 2000US-232685P.

XX 22-SEP-2000; 2000US-234842P.

XX 29-SEP-2000; 2000US-236982P.

XX 05-OCT-2000; 2000US-239057P.

XX 13-OCT-2000; 2000US-240540P.

XX 18-OCT-2000; 2000US-241700P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lee EA, Yue H, Lai PC, Walla NK, Baughn MR, Warren BA, Lee S,

PI Sanjanaala MS, Yao MG, Ramkumar J, Thornton M, Gandhi AR, Naini A,

PI Policky JL, Elliott VS, Arvizu C, Raumann BE, Bruns CM, Naini A,

PI Hatfield AA, Nguyen DB, Xu Y, Lu DAM, Ison CH, Griffin JA,

PI Reddy RM, Burford N;

XX WPI: 2002-393948/42.

XX N-PSDB: AAD36313.

XX Polypeptides of human transporters and ion channels, useful for

PT diagnosing, treating or preventing transport, neurological, muscle,

PT immunological and cell proliferative disorders

XX Claim 1; Page 163-165; 204pp; English.

XX The invention relates to human transporters and ion channels (TRICH)

CC and their corresponding nucleic acid sequences. TRICH is useful for

CC screening an agonist/antagonist that modulates its activity. TRICH is

CC useful as an immunogen for preparing antibodies which are useful for

CC diagnosing a condition of disease associated with its expression in a

CC subject, and for detecting and purifying it from a sample. TRICH DNA

CC is useful as probe or a primer for assessing toxicity of a test

CC compound. Composition comprising TRICH or its agonist is useful for

CC treating a disease or condition associated with decreased expression

CC of functional TRICH and composition comprising TRICH antagonist is

CC useful for treating a disease or condition associated with TRICH

CC overexpression of TRICH. TRICH sequence is used in the diagnosis and

CC treatment of transport disorder e.g. diabetes mellitus, angina,

CC Alzheimer's disease; neurological disorder e.g. epilepsy, stroke,

CC Huntington's disease, bacterial and viral meningitis, muscle disorder

CC e.g. myocarditis, infectious myositis, arrhythmias, asthma,

CC immunological disorder e.g. acquired immunodeficiency syndrome (AIDS),

CC allergies, atherosclerosis; and cell proliferative disorders e.g.

CC cirrhosis, hepatitis, psoriasis and cancers. TRICH DNA is used in

CC gene therapy. TRICH DNA is useful for creating knockin humanised

CC animals (pigs) or transgenic animals (mice or rats) to model human

CC disease. The present sequence is human TRICH protein.

XX

XX Sequence 791 AA:

XX

XX Query Match 30.1%; Score 50; DB 23; Length 791;

XX Best Local Similarity 53.8%; Pred. No. 58;

XX Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

XX

XX 4 SGSSFVGQLFGQHPDYFLMEPAWHV 29

XX :||| ||||| |||||

XX Db 37 AGSIF--SFGILPLVVF-WRPAMHV 58

XX

XX RESULT 33

XX ID ABB57815

XX ABB57815 standard; Protein: 4097 AA.

XX AC ABB57815;

XX

XX 26-MAR-2002 (first entry)

XX

XX Drosophila melanogaster polypeptide SEQ ID NO 237.

XX

XX Drosophila: developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-194637P.

XX 11-JUL-2000; 2000US-0644150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI: 2001-656860/75.

PI N-PSDB: ABL01918.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions

XX Disclosure; SEQ ID NO 237; 21pp + Sequence Listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

CC sequences (AB101840-AB116175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 4097 AA;

XX 26-JUN-2001 (first entry)
 DT Human protein sequence SEQ ID NO:12564.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KM
 XX Homo sapiens.
 OS
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PR 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 12564; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 XX
 SQ Sequence 265 AA;
 XX
 XX Query Match 29.5%; Score 49; DB 22; Length 265;
 XX Best Local Similarity 41.4%; Pred. No. 23;
 XX Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 XX
 QY 1 SWR---SGSSFFVGQLFGQHPDVFYIMEPA 26
 XX :|: ||: : || | | |||: ||
 DB 219 AWKDVTSGNNAKIVQLPGGH---FYLLDPA 244
 XX
 RESULT 36
 ABB97518
 ID ABB97518 standard; Protein: 265 AA.
 XX

AC ABB97518;
 XX
 XX 27-JUN-2002 (first entry)
 DT Novel human protein SEQ ID NO: 786.
 XX
 DE Human; antinaemic; vulnerrary; antiinflammatory; immunomodulator;
 KM antileptility; cerebroprotective; cyostatic; rheumatic; gene therapy;
 KM neuroprotective; antiparinsonian; protein therapy; EST;
 KM expressed sequence tag.
 XX
 OS Homo sapiens.
 XX
 PN W0200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PE 10-SEP-2001; 2001MO-US26015.
 XX
 PR 11-SEP-2000; 2000US-0659671.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI: 2002-292408/33.
 DR N-PSDB; ABB32704.
 DR
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PT
 XX
 PS Claim 20; SEQ ID NO 786; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haemotopoesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. Rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 CC
 XX
 SQ Sequence 265 AA;
 XX
 XX Query Match 29.5%; Score 49; DB 23; Length 265;
 XX Best Local Similarity 41.4%; Pred. No. 23;
 XX Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 XX
 QY 1 SWR---SGSSFFVGQLFGQHPDVFYIMEPA 26
 XX :|: ||: : || | | |||: ||
 DB 219 AWKDVTSGNNAKIVQLPGGH---FYLLDPA 244
 XX
 RESULT 37
 AAU75501
 ID AAU75501 standard; Protein: 265 AA.
 XX
 AC AAU75501;
 XX
 XX 23-APR-2002 (first entry)
 DT Human s-acyl fatty acid synthase thioesterase-like enzyme #1.
 DE
 XX S-acyl fatty acid synthase thioesterase-like; SFST-like; enzyme;
 KM cardiomy; anorectic; vasotropic; extracellular matrix degradation;
 KM cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;
 KM wasting disorder; appetite suppression; appetite enhancement; bulimia;
 KM diabetes; congestive heart failure; myocardial infarction; human;
 KM ischaemic disease; atrial arrhythmia; ventricular arrhythmia;
 KM hypertensive vascular disease; peripheral vascular disease.
 KM

XX OS Homo sapiens.
 XX PN WO200200855-A2.
 XX PD 03-JAN-2002.
 XX PF 26-JUN-2001; 2001WO-EP07297.
 XX PR 26-JUN-2000; 2000US-214012P.
 XX PR 14-DEC-2000; 2000US-255148P.
 XX PA (FARB) BAYER AG.
 XX PI xiao Y;
 XX DR WPI: 2002-130886/17.
 XX DR N-PSDB; ABR13701.
 PT New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
 PT useful for identifying modulators of enzyme activity for treating
 PT cardiovascular disease, diabetes, obesity and hyperlipidaemia -
 XX PS Claim 1; Fig 2; 123pp: English.
 CC The invention describes a purified human S-acyl fatty acid synthase
 CC thioesterase (SFS1)-like enzyme (I). (I) is useful for screening for
 CC agents which decrease or regulate the activity of an SFS1-like enzyme and
 CC agents which decrease extracellular matrix degradation. The
 CC polynucleotide (II) is useful for detecting a polynucleotide which
 CC encodes (I) in a biological sample by formation of a hybridisation
 CC complex. A reagent modulating the activity of SFS1-like enzyme or an
 CC antibody can also be used to detect the polypeptide or polynucleotide in
 CC a biological sample. A reagent modulating the activity of (I) or (II) is
 CC also useful for treating a SFS1-like enzyme dysfunction related disease
 CC condition such as cardiovascular disease, hyperlipidaemia, obesity,
 CC anorexia, cachexia, wasting disorders, appetite suppression, appetite
 CC enhancement, bulimia or diabetes. The cardiovascular diseases treated by
 CC the above mentioned methods and reagents include congestive heart
 CC failure, myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases, and
 CC peripheral vascular diseases. (I) is useful in diagnostic assays for
 CC detecting diseases and abnormalities related to presence of mutations in
 CC the nucleic acid sequences which encode the enzyme. This is the amino
 CC acid sequence of a human s-acyl fatty acid synthase thioesterase-like
 CC enzyme, described in the method of the invention.
 XX SQ Sequence 265 AA:
 Query Match 29.5%; Score 49; DB 23; Length 265;
 Best Local Similarity 41.4%; Pred. No. 23;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 OY 1 SMR---SGSSFYQLFGQHPDYTYLMEPA 26
 DB 219 AMKDVTSNGNAKIYQLPGH---FYLLDPA 244
 RESULT 38
 AAU75504
 ID AAU75504 standard; Protein; 265 AA.
 AC AAU75504;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S-acyl fatty acid synthase thioesterase-like enzyme (short).
 XX
 KM S-acyl fatty acid synthase thioesterase-like; SFS1-like; enzyme;
 KM cardiant; anorectic; vasotropic; extracellular matrix degradation;
 KM cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;
 KM wasting disorder; appetite suppression; appetite enhancement; bulimia;
 KM diabetes; congestive heart failure; myocardial infarction; human;

KM ischemic disease; atrial arrhythmia; ventricular arrhythmia;
 KM hypertensive vascular disease; peripheral vascular disease.
 XX OS Homo sapiens.
 XX PN WO200200855-A2.
 XX PD 03-JAN-2002.
 XX PF 26-JUN-2001; 2001WO-EP07297.
 XX PR 26-JUN-2000; 2000US-214012P.
 XX PR 14-DEC-2000; 2000US-255148P.
 XX PA (FARB) BAYER AG.
 XX PI xiao Y;
 XX DR WPI: 2002-130886/17.
 XX DR N-PSDB; ABR13701.
 PT New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
 PT useful for identifying modulators of enzyme activity for treating
 PT cardiovascular disease, diabetes, obesity and hyperlipidaemia -
 XX PS Disclosure: Fig 14; 123pp: English.
 CC The invention describes a purified human S-acyl fatty acid synthase
 CC thioesterase (SFS1)-like enzyme (I). (I) is useful for screening for
 CC agents which decrease or regulate the activity of an SFS1-like enzyme and
 CC agents which decrease extracellular matrix degradation. The
 CC polynucleotide (II) is useful for detecting a polynucleotide which
 CC encodes (I) in a biological sample by formation of a hybridisation
 CC complex. A reagent modulating the activity of SFS1-like enzyme or an
 CC antibody can also be used to detect the polypeptide or polynucleotide in
 CC a biological sample. A reagent modulating the activity of (I) or (II) is
 CC also useful for treating a SFS1-like enzyme dysfunction related disease
 CC condition such as cardiovascular disease, hyperlipidaemia, obesity,
 CC anorexia, cachexia, wasting disorders, appetite suppression, appetite
 CC enhancement, bulimia or diabetes. The cardiovascular diseases treated by
 CC the above mentioned methods and reagents include congestive heart
 CC failure, myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases, and
 CC peripheral vascular diseases. (I) is useful in diagnostic assays for
 CC detecting diseases and abnormalities related to presence of mutations in
 CC the nucleic acid sequences which encode the enzyme. This sequence encodes
 CC a human s-acyl fatty acid synthase thioesterase-like enzyme, described
 CC in the method of the invention.
 XX SQ Sequence 265 AA:
 Query Match 29.5%; Score 49; DB 23; Length 265;
 Best Local Similarity 41.4%; Pred. No. 23;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 OY 1 SMR---SGSSFYQLFGQHPDYTYLMEPA 26
 DB 219 AMKDVTSNGNAKIYQLPGH---FYLLDPA 244
 RESULT 39
 AAB93499
 ID AAB93499 standard; Protein; 318 AA.
 AC AAB93499;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12816.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KM Homo sapiens.

XX EP1074617-A2.
PD 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WFI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 12816; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92246 to
CC AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 318 AA;

Query Match 29.5%; Score 49; DB 22; Length 318;
Best Local Similarity 41.4%; Pred. No. 29;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2

QY 1 SMR---SGSSFFGOLFQGHDFVYLMPEA 26
:|: ||::||| | ||::||
Db 272 AKKDVYSGNAKITQLPGH---FYLLDPA 297

RESULT 40
ID AAU75505
AAU75505 standard; Protein; 318 AA.
XX AAU75505;
AC
XX 23-APR-2002 (first entry)
DT
XX
XX S-acyl fatty acid synthase thioesterase-like enzyme (long).
DE
XX
XX S-acyl fatty acid synthase thioesterase-like, SRSF-like; enzyme;
KW cardian; anorectic; vasotrophic; extracellular matrix degradation;
KW cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;

KM	wasting disorder; appetite suppression; appetite enhancement; bulimia;
KM	diabetes; congestive heart failure; myocardial infarction; human;
KM	ischemic disease; atrial arrhythmia; ventricular arrhythmia;
KM	hypertensive vascular disease; peripheral vascular disease.
XX	
OS	Homo sapiens.
PN	W0200200855-A2.
PD	03-JAN-2002.
PF	26-JUN-2001; 2001WO-EP07297.
PR	26-JUN-2000; 2000US-214012P.
PR	14-DEC-2000; 2000US-255148P.
XX	
PA	(FARB) BAYER AG.
PI	Xiao Y;
XX	
DR	WPI: 2002-130886/17.
DR	N-PDSB; ABK13711.
PT	New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
PT	useful for identifying modulators of enzyme activity for treating
PT	cardiovascular disease, diabetes, obesity and hyperlipidemia -
XX	
PS	Disclosure: Fig 16; 123pp; English.
XX	
CC	The invention describes a purified human S-acyl fatty acid synthase
CC	thioesterase (SSTT)-like enzyme (I). (I) is useful for screening for
CC	agents which decrease or regulate the activity of an SSTT-like enzyme and
CC	SSTT agents which decrease extracellular matrix degradation. The
CC	polynucleotide (II) is useful for detecting a polynucleotide which
CC	encodes (I) in a biological sample by formation of a hybridisation
CC	complex. A reagent modulating the activity or SSTT-like enzyme or an
CC	antibody can also be used to detect the polypeptide or polynucleotide in
CC	a biological sample. A reagent modulating the activity of (I) or (II) is
CC	also useful for treating a SSTT-like enzyme dysfunction related disease
CC	condition such as cardiovascular disease, hyperlipidemia, obesity,
CC	anorexia, cachexia, wasting disorders, appetite suppression, appetite
CC	enhancement, bulimia or diabetes. The cardiovascular diseases treated by
CC	the above mentioned methods and reagents include congestive heart
CC	failure, myocardial infarction, ischemic diseases of the heart, atrial
CC	and ventricular arrhythmia, hypertensive vascular diseases, and
CC	peripheral vascular diseases. (I) is useful in diagnostic assays for
CC	detecting diseases and abnormalities related to presence of mutations in
CC	the nucleic acid sequences which encode the enzyme. This sequence encodes
CC	a human s-acyl fatty acid synthase thioesterase-like enzyme, described
CC	in the method of the invention.
XX	
SO	Sequence 318 AA;
QY	Query Match 29.5%; Score 49; DB 23; Length 318; Best Local Similarity 41.4%; Pred. No. 29; Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
DB	1 SWR---SGSSFYGQLFGQHDPVFIYMERA 26 ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: 272 AMKDYTSGNAKTYQLPGH--FYLLDPA 297
RESULT 41	
ID	AAZ27357 standard; protein; 931 AA.
AC	AAZ27357;
DT	15-NOV-1999 (first entry)
DE	Group B Streptococcus (GBS) antigen (clone 4).
KM	Group B Streptococcus; GBS; antigen; vaccine; Streptococcus infection;

XX MPI: 2002-352536/38.
DR N-PSDB: ABN67342.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1: Page 3409; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 1252 AA;
SQ
Query Match 29.2%; Score 48.5; DB 23; Length 1252;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 10; Conservative 5; Mismatches 4; Indels 1; Gaps 1.
QY 1 SWRSSSEFYQG-LFGQHPV 19
:| ||||| ||:| ||
DB 304 TWPDSNPFYNGLYGRYDV 323
RESULT 44
ABP29749
ID ABP29749 standard; Protein; 1252 AA.
XX
AC ABP29749;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 8674.
XX
KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
XX WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PE 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margarit Ros YL, Grandi G, Fraser C;
PI Tettelin H;
XX
DR MPI: 2002-352536/38.
DR N-PSDB: ABN70380.

XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3982; 4525bp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN65044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
CC
CC Sequence 1252 AA;
SQ
OY 1 SWRSGSSFEYVGQ-LFGQHPDV 19
DB 304 TWPDGSMFNYOGLYGRYIDV 323
RESULT 45
ABB49146
ID ABB49146 standard; Protein; 240 AA.
XX
XX ABB49146;
AC
XX
XX 05-FEB-2002 (first entry)
DE Listeria monocytogenes protein #1850.
XX
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KM vitamin B12; bacterial infection; disease.
XX
XX Listeria monocytogenes.
XX
XX WO200177335-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 11-APR-2001; 2001WO-FR01118.
PF
XX
XX 11-APR-2000; 2000FR-0004629.
PR
XX
XX (INSP) INST PASTEUR.
PA
XX
XX Buchrieser C, Frangoul L, Couve E, Rusniok C, Fsihi H, Dehoux P,
PI Dussurgeat O, Chelouani F, Nedjati H, Glaser P, Kunst F, Cossart P,
PI Darsels J, Goebel W, Kretz J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Doman E, Hahn T, Berche P, Charbit A, Durant L;
PI Perez-Izaz J, Requero F, Garcia Del Portillo F, Gomez-Lopez N,
PI Madeno A, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX
XX Rose M, Voss H;
XX
XX WPI: 2002-010914/01.
DR
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides -
 XX
 PS Claim 6; SEQ ID No 1851; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC monocytes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC monocytes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L.* monocytes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L.* monocytes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC monocytes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 240 AA;
 OY
 DB 8 FVQGLFGQHPDVFYLMPE 25
 DB 143 FIARALQHAHFIFLDEP 160
 XX
 RESULT 46
 ABB53040
 ID ABB53040 standard; Protein; 288 AA.
 XX
 AC ABB53040;
 XX
 DT 11-FEB-2002 (first entry)
 XX
 DE *Escherichia coli* polypeptide SEQ ID NO 1449.
 XX
 KM *Escherichia coli*; B2/D+A-; antiinflammatory; antibacterial;
 KM immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KM systemic infection; non-diarrhoeal infection; septicemia;
 KM pyelonephritis; antibiotic resistance.
 XX
 OS *Escherichia coli*.
 XX
 PN WO200166572-A2.
 PD
 PD 13-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-EP03445.
 XX
 PR 10-MAR-2000; 2000FR-0003145.
 PR 02-FEB-2001; 2001FR-0001449.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Bingen E, Bonaccorsi S, Clermont O, Nassif X, Tinsley C;
 XX
 DR WPI; 2001-550253/61.
 XX
 PT A library of DNA fragments of *Escherichia coli* strains for the
 PT phylogenetic determination of a given strain comprises polynucleotides of
 PT nature B2/D+ A- -
 XX
 PS Example 6; Fig 6; 646pp; English.
 XX

CC The invention relates to a library of DNA fragments of *Escherichia coli*
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
 CC and encoded proteins (ABB52459-ABB52519 and ABB52954-ABB53094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal *E. coli*
 CC infections. The polypeptides are useful for determining the phylogenic
 CC group of a given *E. coli* strain. These polypeptides can detect and treat
 CC in undesired development of *E. coli*, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more
 CC frequent use of broad spectrum antibiotics.
 CC
 SQ Sequence 288 AA;
 OY
 DB 18 DVFYLMPEAMH 28
 DB 177 DERFLMEAKWH 187
 XX
 RESULT 47
 AAU33561
 ID AAU33561 standard; Protein; 460 AA.
 XX
 AC AAU33561;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE *Pseudomonas aeruginosa* cellular proliferation protein #5.
 XX
 KM Antisense: prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS *Pseudomonas aeruginosa*.
 XX
 PN WO200170955-A2.
 PD
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS51420.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 5057; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:11 / Search time 46.4 Seconds
(without alignments)
60.084 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166

Sequence: 1 SWRSGSSFFVGOLFQGHDPVFLMEPAMHV 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :
1: PIR1:***
2: PIR2:***
3: PIR3:***
4: PIR4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134	80.7	484	2 UC7350	N-acetylglucosamin
2	134	80.7	486	2 UC7351	N-acetylglucosamin
3	134	80.7	484	2 JE0261	N-acetylglucosamin
4	101	60.8	458	2 AC2012	chondroitin 6-sulf
5	57	34.3	1570	2 G89833	hypothetical prote
6	53	31.9	277	2 G69044	teichoic acid tran
7	53	31.9	390	2 B83472	N-acetylornithine
8	52	31.3	518	2 F97381	hypothetical prote
9	52	31.3	614	2 AE2599	hypothetical prote
10	52	31.3	614	2 AG2164	hypothetical prote
11	50	30.1	1051	2 T48933	hypothetical prote
12	50	30.1	1645	2 A37792	WD repeat domain p
13	50	30.1	1645	2 C83530	spectrin beta-H ch
14	49.5	29.8	634	2 T42993	potassium uptake p
15	49	29.5	4063	2 T23630	probable spectrin
16	49	29.5	4101	2 T23630	hypothetical prote
17	48.5	29.2	449	2 A11305	hypothetical prote
18	48	28.9	240	2 A11305	hypothetical prote
19	48	28.9	240	2 A11305	metal cations ABC
20	48	28.9	245	2 D70670	hypothetical prote
21	48	28.9	338	2 S73902	UDP-glucose 4-epim
22	48	28.9	443	2 T10801	GDP dissociation i
23	48	28.9	460	2 E83617	probable aminotran
24	48	28.9	478	2 F70322	glutamyl-cRNA (in
25	47.5	28.6	518	2 T23120	hypothetical prote
26	47.5	28.6	495	2 D71307	conserved hypothet
27	47	28.3	176	2 G73781	hypothetical prote
28	47	28.3	284	2 A83726	phosphatidate cycl
29	47	28.3	288	2 F84293	hypothetical prote

30	47	28.3	296	2 H98338	regulatory protein
31	47	28.3	296	2 AB2944	transcription regu
32	47	28.3	349	2 F86649	dtbp-glucose 4,6-d
33	47	28.3	363	2 T20608	hypothetical prote
34	47	28.3	381	2 T15140	hypothetical prote
35	47	28.3	444	2 D98317	hypothetical prote
36	47	28.3	444	2 AH2965	probable pyridoxal
37	47	28.3	465	2 S76464	hypothetical prote
38	47	28.3	465	2 S51613	hypothetical prote
39	47	28.3	701	2 C83052	cyclin homolog CLN
40	47	28.3	1340	2 D96521	polypurbonucleotide
41	47	28.3	3013	2 AB0480	protein F21D18.16
42	46.5	28.0	251	1 ESDKTM	probable invasion y
43	46.5	28.0	251	1 I50520	oleoyl-lacetyl-carri
44	46.5	28.0	542	2 A28550	S-acyl fatty acid
45	46.5	28.0	634	2 F82623	cyclohexanone mono
46	46.5	28.0	820	2 H86246	potassium uptake p
47	46	27.7	175	2 T08948	hypothetical prote
48	46	27.7	234	2 T46754	Adcc protein limpo
49	46	27.7	234	2 H95253	hypothetical prote
50	46	27.7	234	2 H95253	zinc ABC transport
51	46	27.7	276	2 S57627	avirulence protein
52	46	27.7	291	2 F83504	probable transcrip
53	46	27.7	364	2 C75597	thymidine diphosph
54	46	27.7	426	2 T48424	hypothetical prote
55	46	27.7	557	2 E86850	hypothetical prote
56	46	27.7	764	2 D83870	ATP-dependent RNA
57	46	27.7	1331	2 T18310	receptor-adenylate
58	45.5	27.4	380	2 T19280	hypothetical prote
59	45.5	27.4	423	1 XNCHDM	aspartate transami
60	45.5	27.4	430	1 XNRTDM	aspartate transami
61	45.5	27.4	433	2 S01174	aspartate transami
62	45.5	27.4	519	2 D65178	kup protein - Esch
63	45.5	27.4	622	2 A49345	K(+)-uptake protei
64	45.5	27.4	622	2 AG0952	membrane transport
65	45.5	27.4	809	2 B96314	probable quinolate d
66	45.5	27.4	809	2 A12968	hypothetical prote
67	45.5	27.4	844	2 S75547	ethylene response
68	45	27.1	214	2 S77391	hypothetical prote
69	45	27.1	218	2 E87552	conserved hypotnet
70	45	27.1	225	2 G83244	probable two-compo
71	45	27.1	225	2 AP2049	ATP-binding protei
72	45	27.1	313	2 D95866	probable transcrip
73	45	27.1	338	2 E72334	conserved hypotnet
74	45	27.1	362	2 A49316	chorismate synthas
75	45	27.1	370	2 G83219	probable dihydrol
76	45	27.1	421	2 C96806	unknown protein T5
77	45	27.1	465	2 A82438	amino transferase,
78	45	27.1	622	2 S65539	GABA/beta-alanine
79	45	27.1	644	1 FGHUA	fibrinogen alpha c
80	45	27.1	648	2 T51220	hypothetical prote
81	45	27.1	669	2 T51220	hypothetical prote
82	45	27.1	808	2 T15562	hypothetical prote
83	45	27.1	866	2 D44234	hypothetical prote
84	45	27.1	884	2 G66684	fibrinogen alpha c
85	45	27.1	1279	2 A12067	prophage pil prote
86	45	27.1	1654	2 A12067	two-component sens
87	45	27.1	1691	1 D54689	protein-tyrosine-p
88	45	27.1	1894	2 C54689	protein-tyrosine-p
89	45	27.1	1912	2 A56178	protein-tyrosine-p
90	44.5	26.8	2505	1 XYRTFA	enoyl-lacetyl-carri
91	44.5	26.8	401	2 F87196	probable membrane
92	44.5	26.8	430	1 XNHUIM	aspartate transami
93	44.5	26.8	430	2 AB3237	aspartate transami
94	44.5	26.8	430	2 S55960	aspartate transami
95	44.5	26.8	469	2 T01128	hypothetical prote
96	44.5	26.8	483	2 G75392	glycosyl hydrolase
97	44.5	26.8	687	2 T27683	hypothetical prote
98	44.5	26.8	1266	2 G97477	hypothetical prote
99	44.5	26.8	1266	2 A52695	conserved hypotnet
100	44	26.5	1396	1 VCBE40	major capsid prote
101	44	26.5	207	2 AG0753	colanic acid capsu
102	44	26.5	218	2 E84840	hypothetical prote
	44	26.5	254	2 S75676	probable phosphoes

103	44	26.5	312	2	T09443	zinc metalloprote
104	44	26.5	312	2	E82276	hypothetical prote
105	44	26.5	312	2	H91237	hypothetical prote
106	44	26.5	312	2	D86085	hypothetical prote
107	44	26.5	312	2	I78664	hypothetical 34.1k
108	44	26.5	326	2	S77304	beta transducin-ii
109	44	26.5	362	2	AR0157	probable periplasm
110	44	26.5	364	2	E75012	hypothetical prote
111	44	26.5	367	2	H75440	alanine racemase -
112	44	26.5	405	1	H69424	probable iron(III)
113	44	26.5	413	2	S43527	gene P protein - p
114	44	26.5	439	2	E98139	hypothetical prote
115	44	26.5	439	2	AP3148	hypothetical prote
116	44	26.5	445	2	T02032	GDP dissociation i
117	44	26.5	457	2	B80008	hypothetical prote
118	44	26.5	460	2	D70435	hypothetical prote
119	44	26.5	483	2	F90777	replication protei
120	44	26.5	496	2	A49930	carb protein homol
121	44	26.5	609	2	F64045	excinuclease ABC c
122	44	26.5	627	2	B44409	gamma-aminobutyric
123	44	26.5	627	2	JH0695	gamma-aminobutyric
124	44	26.5	633	2	G95385	kmp2 Potassium upt
125	44	26.5	638	2	H83905	hypothetical prote
126	44	26.5	812	2	E71972	pyruvate, water di
127	44	26.5	812	2	A64535	pyruvate, water di
128	44	26.5	837	2	H72802	minor tail subunit
129	44	26.5	2105	1	A44059	genome polyprotein
130	44	26.5	2504	1	A57768	enoyl-lacyl-carrie
131	44	26.5	2509	2	G01880	fatty-acid synthas
132	43.5	26.2	157	2	C70477	conserved hypotet
133	43.5	26.2	320	2	C83628	probable oxidoredu
134	43.5	26.2	328	2	E85651	hypothetical prote
135	43.5	26.2	328	2	C64845	probable membrane
136	43.5	26.2	328	2	B80791	hypothetical prote
137	43.5	26.2	552	2	B82633	ubiquinone biosynt
138	43.5	26.2	564	2	AR2328	AMP-binding protei
139	43.5	26.2	1111	1	A33288	mposin heavy chain
140	43	25.9	146	2	A44588	globin - waved whe
141	43	25.9	147	2	AF1905	hypothetical prote
142	43	25.9	150	2	S01291	hypothetical prote
143	43	25.9	153	2	A55139	myoglobin, body wa
144	43	25.9	162	2	A31281	hypothetical prote
145	43	25.9	162	2	A11643	hypothetical prote
146	43	25.9	165	2	C70939	hypothetical prote
147	43	25.9	245	2	B86889	zinc ABC transport
148	43	25.9	306	2	D70924	probable cdsa prot
149	43	25.9	311	1	JC1191	osteopontin precur
150	43	25.9	352	2	A39042	fatty-acid synthas
151	43	25.9	355	2	T07160	glutamate-ammonia
152	43	25.9	360	2	B70863	hypothetical prote
153	43	25.9	373	2	G75338	conserved hypotet
154	43	25.9	375	2	E82292	chorismate mutase/
155	43	25.9	404	2	B71247	hypothetical prote
156	43	25.9	404	2	S50648	hypothetical prote
157	43	25.9	413	2	S51667	serine/threonine k
158	43	25.9	427	2	AF2222	nitrate-binding pr
159	43	25.9	444	2	E26421	shufflon C - Esche
160	43	25.9	445	2	S27492	hypothetical prote
161	43	25.9	445	2	T02030	GDP dissociation i
162	43	25.9	461	1	DCMUO	ornithine decarbox
163	43	25.9	473	1	A42391	Ca2+-transporter
164	43	25.9	474	2	AD3312	hypothetical ATP-b
165	43	25.9	493	1	S39532	aldehyde dehydroge
166	43	25.9	511	2	S49151	maturase mark, int
167	43	25.9	572	2	T37128	hypothetical prote
168	43	25.9	608	2	B87575	ABC transporter, A
169	43	25.9	694	2	B91161	4-alpha-glucanotra
170	43	25.9	694	2	A86007	4-alpha-glucanotra
171	43	25.9	733	2	S44254	alpha-galactosidas
172	43	25.9	1069	2	AR1930	hypothetical prote
173	43	25.9	1168	2	H88816	transcription-repa
174	43	25.9	1247	2	A33812	interphotoreceptor
175	43	25.9	1273	2	S58782	SEC31 protein - ye
176	43	25.9	176	2	T29090	surface layer-asso
177	43	25.9	177	2	G84922	hypothetical prote
178	43	25.9	178	2	A54849	collagen alpha 1(V
179	43	25.9	179	2	T08839	polyprotein - marm
180	43	25.9	3421	1	WZBBE6	367k tegument prot
181	43	25.9	3660	1	S02041	dystrrophin, muscle
182	42.5	25.6	171	2	A99190	hypothetical prote
183	42.5	25.6	171	2	AH3096	conserved hypotet
184	42.5	25.6	172	2	S24398	stellate hypotet
185	42.5	25.6	179	2	AD2007	hypothetical prote
186	42.5	25.6	240	2	S75017	hypothetical prote
187	42.5	25.6	262	2	S75891	hypothetical prote
188	42.5	25.6	305	2	D70432	conserved hypotet
189	42.5	25.6	308	2	F81695	lacyl-carrier-prot
190	42.5	25.6	374	2	AH1825	protoporphyrin ix
191	42.5	25.6	401	2	B26341	aspartate transam
192	42.5	25.6	450	1	ANP6DM	aspartate transam
193	42.5	25.6	527	2	A83453	probable flavin-co
194	42.5	25.6	574	2	S74849	ABC-type transport
195	42.5	25.6	622	2	A98215	low affinity potas
196	42.5	25.6	622	2	B85061	low affinity potas
197	42.5	25.6	833	2	C81835	probable P-type ca
198	42.5	25.6	892	2	T01899	disease resistance
199	42.5	25.6	1069	2	D81826	exodeoxyribonuclea
200	42.5	25.6	1069	2	A81050	exodeoxyribonuclea
201	42.5	25.6	2422	2	T12687	ALR protein homolo
202	42	25.3	174	2	A49181	alpha B-crystallin
203	42	25.3	174	2	S58758	alpha B-crystallin
204	42	25.3	225	2	G75167	hypothetical prote
205	42	25.3	253	2	T31021	hypothetical prote
206	42	25.3	254	2	PD0018	hypothetical prote
207	42	25.3	264	2	F83952	miscellaneous lectin
208	42	25.3	279	2	S65990	phosphatidate cycl
209	42	25.3	280	2	T36376	conserved hypotet
210	42	25.3	290	2	A81775	hypothetical prote
211	42	25.3	290	2	A81199	probable oxidoredu
212	42	25.3	312	2	G87107	prephenate dehydro
213	42	25.3	313	1	FOYDA	probable phosphat
214	42	25.3	339	2	F83026	gag polypeptide -
215	42	25.3	357	2	H84240	conserved hypotet
216	42	25.3	367	1	MHCH	integrase/recombin
217	42	25.3	387	2	D84813	Ig mu chain C regi
218	42	25.3	387	2	S77268	hypothetical prote
219	42	25.3	392	2	F87321	carboxymyosin
220	42	25.3	398	2	S76685	succinylornithine
221	42	25.3	403	2	D75577	hypothetical prote
222	42	25.3	404	2	T32113	flavohemoprotein -
223	42	25.3	405	2	T14311	hypothetical prote
224	42	25.3	418	2	D81740	aspartate transam
225	42	25.3	423	2	S46315	aspartate transam
226	42	25.3	424	2	A87016	hypothetical prote
227	42	25.3	428	2	G89982	probable zinc prot
228	42	25.3	444	2	T01782	hypothetical prote
229	42	25.3	445	2	T44899	GDP dissociation i
230	42	25.3	448	2	D81740	probable proteinas
231	42	25.3	449	2	T48511	conserved hypotet
232	42	25.3	450	2	A44751	aspartate transam
233	42	25.3	451	2	JE0166	carotenoid-binding
234	42	25.3	454	2	AT1606	nitric-oxide reduc
235	42	25.3	503	1	YFB1AC	acetyl-coA carboxy
236	42	25.3	520	2	A13596	phenylalanine-tRNA
237	42	25.3	531	2	AG0181	sugar transport AT
238	42	25.3	571	2	T02213	ABC transporter AT
239	42	25.3	583	2	H69165	NBS-LRR type resis
240	42	25.3	667	2	H98141	hypothetical prote
241	42	25.3	718	1	ALBS66	hypothetical prote
242	42	25.3	718	1	ALBS66	cyclomaltoextrin
243	42	25.3	718	1	ALBS66	cyclomaltoextrin
244	42	25.3	795	2	T52516	hypothetical prote
245	42	25.3	861	2	S77086	hypothetical prote
246	42	25.3	876	2	G89952	DNA polymerase I
247	42	25.3	902	2	A60560	formyltetrahydrofo
248	42	25.3	905	2	T26430	hypothetical prote
249	42	25.3	905	2	T13930	tripeptidyl-peptid

249	42	25.3	1816	2	F83901	hypothetical prote	322	41	24.7	416	2	T02194	probable pectinase
250	41.5	25.0	175	2	F86439	hypothetical prote	323	41	24.7	416	2	B95254	dliD protein (limo
251	41.5	25.0	184	2	A86461	hypothetical prote	324	41	24.7	427	2	B98118	hypothetical prote
252	41.5	25.0	189	1	LEVCTP	thermostable direc	325	41	24.7	439	2	E97079	sugar-binding peri
253	41.5	25.0	250	2	T40937	cdp-diacylglycerol	326	41	24.7	441	2	B84854	hypothetical prote
254	41.5	25.0	304	2	B91034	probable fructokin	327	41	24.7	446	2	G70774	probable atph prot
255	41.5	25.0	305	1	S52161	probable fructokin	328	41	24.7	452	2	T28094	hypothetical prote
256	41.5	25.0	305	2	C83878	D-fructokinase (lm	329	41	24.7	461	2	AH2923	hypothetical prote
257	41.5	25.0	305	2	G75069	ADP-heptose-3S hyd	330	41	24.7	463	2	T45074	Sun protein (limp
258	41.5	25.0	346	2	F82349	ADP-heptose-3S he	331	41	24.7	463	2	H97697	probable phenylala
259	41.5	25.0	348	2	S34494	csca protein - Eug	332	41	24.7	472	2	C42391	probable sun prote
260	41.5	25.0	410	1	IXBE10	alpha trans-induci	333	41	24.7	476	2	E83006	Ca2+-transporting
261	41.5	25.0	437	2	S15306	CDP-4-keto-6-deoxy	334	41	24.7	477	2	T33531	two-component resp
262	41.5	25.0	482	2	C82404	6-phosphogluconate	335	41	24.7	495	2	T05388	hypothetical prote
263	41.5	25.0	485	2	S73655	anthranilate synth	336	41	24.7	496	2	T49698	hypothetical prote
264	41.5	25.0	504	2	A82205	probable deoxyribo	337	41	24.7	502	2	C84400	hypothetical prote
265	41.5	25.0	510	2	JC4208	nitrogenase NifB c	338	41	24.7	514	2	AC3563	phenylalanyl-tRNA
266	41.5	25.0	537	2	T09237	nifb protein - Fra	339	41	24.7	530	2	A90741	ribose transport A
267	41.5	25.0	537	2	S78195	C4-dicarboxylate t	340	41	24.7	530	2	D85591	probable ABC-type
268	41.5	25.0	597	2	B82140	cytochrome-c oxida	341	41	24.7	530	2	D64819	probable ABC-type
269	41.5	25.0	620	2	F83976	cytochrome-c oxida	342	41	24.7	531	2	AD0602	probable ABC-type
270	41.5	25.0	622	2	E86609	cytochrome-c oxida	343	41	24.7	536	2	S40717	ABC transporter AT
271	41.5	25.0	635	2	D45335	cytochrome-c oxida	344	41	24.7	540	2	B45665	hypothetical prote
272	41.5	25.0	642	2	T51421	L-aspartate oxida	345	41	24.7	550	2	A48053	adult-specific 6L
273	41.5	25.0	706	2	S62501	hypothetical prote	346	41	24.7	553	2	C71257	hypothetical prote
274	41.5	25.0	764	2	D84847	probable receptor-	347	41	24.7	589	2	F86202	phenylalanine-tRNA
275	41.5	25.0	802	2	G87578	Tom-dependent rec	348	41	24.7	622	2	A40144	hypothetical prote
276	41.5	25.0	823	2	D81129	calion transport A	349	41	24.7	636	2	C81128	prolactin receptor
277	41.5	25.0	944	2	T47246	chitin synthase (E	350	41	24.7	636	2	A81893	probable ABC trans
278	41.5	25.0	993	2	C82687	glycine decarboxyl	351	41	24.7	641	2	A24075	ABC transporter, A
279	41.5	25.0	1215	2	T32734	myosin-1A - Acanth	352	41	24.7	645	2	G01205	lipase precursor -
280	41.5	25.0	1400	2	T22644	hypothetical prote	353	41	24.7	669	2	E71127	tyl protein - huma
281	41.5	25.0	2396	2	T13714	kakapo gene protei	354	41	24.7	671	2	H86702	hypothetical prote
282	41.5	25.0	2535	2	AC0304	probable hemolysin	355	41	24.7	843	1	A27340	potassium uptake p
283	41.5	25.0	2833	2	T23064	hypothetical prote	356	41	24.7	864	1	JC1422	complement C7 prec
284	41.5	25.0	2823	2	F87908	protein T22A3.8 [1	357	41	24.7	884	2	T40690	nitrate reductase
285	41.5	25.0	3102	2	T43291	laminin alpha chai	358	41	24.7	891	2	B48642	hypothetical prote
286	41.5	25.0	152	2	G71955	hypothetical prote	359	41	24.7	925	2	T01384	acornate hydratase
287	41.5	25.0	177	2	T37137	probable acetyltra	360	41	24.7	934	2	161714	hypothetical prote
288	41.5	25.0	178	2	G02294	plasma membrane ca	361	41	24.7	990	2	T03784	co-repressor prote
289	41.5	25.0	179	2	S75860	hypothetical prote	362	41	24.7	1016	1	J50428	probable receptor
290	41.5	25.0	183	2	F69049	conserved hypotet	363	41	24.7	1060	2	F88710	NAD ADP-ribosyltra
291	41.5	25.0	196	1	A04PNV	aequorin precursor	364	41	24.7	1079	2	T30956	protein C015.4 [1
292	41.5	25.0	196	2	A26623	aequorin-1 precurs	365	41	24.7	1403	2	H67131	hypothetical prote
293	41.5	25.0	205	2	AD0774	probable exported	366	41	24.7	1449	2	B81963	uncharacterized, p
294	41.5	25.0	224	2	T01185	protein kinase (EC	367	41	24.7	2064	2	T13707	194-specific serin
295	41.5	25.0	228	2	C81243	conserved hypotet	368	41	24.7	2242	2	A57541	still life protein
296	41.5	25.0	230	2	JC1483	deoxyribonuclease	369	41	24.7	108	2	S21420	pyrimidine synthe
297	41.5	25.0	234	2	H89920	conserved hypotet	370	41	24.4	156	2	S67248	general stress pro
298	41.5	25.0	241	2	H81278	phosphatidate cyti	371	41	24.4	185	2	JC7369	hypothetical prote
299	41.5	25.0	247	1	D24706	nodulation protein	372	41	24.4	242	2	A10659	phenylated Rab acc
300	41.5	25.0	247	2	H95319	phosphadenylyl-su	373	41	24.4	252	2	A12088	conserved hypotet
301	41.5	25.0	254	2	D84559	diethylsulfoxide	374	41	24.4	266	2	A51186	acyl-lactyl-carrier
302	41.5	25.0	279	2	F64109	prolactin receptor	375	41	24.4	269	2	T20670	hypothetical prote
303	41.5	25.0	280	2	G81651	prolactin receptor	376	41	24.4	270	2	H64433	hypothetical prote
304	41.5	25.0	288	2	B58405	probable GDSL-moti	377	41	24.4	280	2	E82059	phosphatidate cyti
305	41.5	25.0	301	2	H82566	hypothetical prote	378	41	24.4	286	2	T21545	hypothetical prote
306	41.5	25.0	303	2	S51460	hypothetical prote	379	41	24.4	301	2	G83352	transcription regu
307	41.5	25.0	312	2	F83955	hypothetical prote	380	41	24.4	322	2	T50557	lactate dehydrogen
308	41.5	25.0	314	2	F85950	hypothetical prote	381	41	24.4	332	2	S49054	4,6-de
309	41.5	25.0	317	2	F85650	hypothetical prote	382	41	24.4	333	2	A47160	L-lactate dehydrog
310	41.5	25.0	330	2	C95844	hypothetical prote	383	41	24.4	335	2	G82247	cysteine synthase/
311	41.5	25.0	330	2	D87068	probable sugar ABC	384	41	24.4	335	2	G83409	hypothetical prote
312	41.5	25.0	340	2	A82218	transcription regu	385	41	24.4	339	2	AE2293	hypothetical prote
313	41.5	25.0	349	2	T01417	hypothetical prote	386	41	24.4	389	2	E83409	hypothetical prote
314	41.5	25.0	356	2	H70882	hypothetical prote	387	41	24.4	399	2	A84275	hypothetical prote
315	41.5	25.0	369	2	A50451	maltoase/maltodextr	388	41	24.4	435	2	A84824	hypothetical prote
316	41.5	25.0	376	2	A59405	prolactin receptor	389	41	24.4	437	2	T04444	cytochrome p450 -
317	41.5	25.0	383	2	C69442	succinyl-CoA synth	390	41	24.4	485	2	C70468	probable nematode-
318	41.5	25.0	389	2	A87602	conserved hypotet	391	41	24.4	496	2	G89824	cytochrome-c oxida
319	41.5	25.0	400	2	AF0291	probable drug resi	392	41	24.4	499	2	A38891	hypothetical prote
320	41.5	25.0	401	2	T39185	probable polysacch	393	41	24.4	512	2	S28663	cytochrome-c oxida
321	41.5	25.0	411	2	T15705	hypothetical prote	394	41	24.4	528	2	JC4814	Na+/H+-exchanging

395	40.5	24.4	540	2	E69861	ABC transporter (A	468	40	24.1	448	2	C97757	hypothetical prote
396	40.5	24.4	549	2	A87347	hypothetical prote	469	40	24.1	454	1	A26955	alkaline serine pr
397	40.5	24.4	552	2	AC3445	cytochrome-c oxida	470	40	24.1	461	1	DCMSO	ornithine decarbox
398	40.5	24.4	556	2	C87472	steroid monooxygen	471	40	24.1	461	1	I55356	ornithine decarbox
399	40.5	24.4	600	2	T48433	hypothetical prote	472	40	24.1	461	2	I56477	ornithine decarbox
400	40.5	24.4	616	2	JX0140	cytochrome-c oxida	473	40	24.1	462	2	T50164	probable phosphor
401	40.5	24.4	642	2	A11837	cyclomaltodextrin	474	40	24.1	467	2	AE3142	hypothetical prote
402	40.5	24.4	673	2	T47006	hypothetical prote	475	40	24.1	468	2	A69611	cytochrome bd ubiq
403	40.5	24.4	673	2	AC0238	heamin storage sys	476	40	24.1	468	2	AD1790	cytochrome D ubiq
404	40.5	24.4	729	2	F86308	similar to disease	477	40	24.1	468	2	AE1414	cytochrome D ubiq
405	40.5	24.4	736	2	T40080	probable ABC trans	478	40	24.1	471	2	PS0154	125k surface antig
406	40.5	24.4	749	2	T43370	oligosaccharyltran	479	40	24.1	473	2	S22621	phosphomannomuta
407	40.5	24.4	752	2	T39338	oligosaccharyl tra	480	40	24.1	477	2	E91136	D-alanyl-D-alanine
408	40.5	24.4	752	2	S56146	GCN20 protein - ye	481	40	24.1	477	2	H85981	D-alanyl-D-alanine
409	40.5	24.4	763	2	A47563	glucose-6-phosphat	482	40	24.1	477	2	A54535	serine-type D-Ala-
410	40.5	24.4	831	2	S62751	probable DNA-direc	483	40	24.1	484	2	E84765	hypothetical prote
411	40.5	24.4	878	1	T01734	hypothetical prote	484	40	24.1	485	2	E70558	probable apcC prot
412	40.5	24.4	1180	2	NCEKX5	exodeoxyribonuclea	485	40	24.1	485	2	S60175	regulatory protein
413	40.5	24.4	1180	2	E91088	DNA helicase RecB	486	40	24.1	497	2	AF0505	probable secreted
414	40.5	24.4	1180	2	G85933	hypothetical prote	487	40	24.1	501	2	T36051	probable cytochrom
415	40.5	24.4	88	2	E82754	hypothetical prote	488	40	24.1	509	2	D82212	conserved hypothet
416	40.5	24.1	122	2	E95414	hypothetical prote	489	40	24.1	520	2	G98145	probable aminotran
417	40.5	24.1	166	2	A75426	hypothetical prote	490	40	24.1	520	2	C81424	cytochrome bd oxid
418	40.5	24.1	189	2	T33747	hypothetical prote	491	40	24.1	529	2	S76831	hypothetical prote
419	40.5	24.1	200	2	B66700	transposon-related	492	40	24.1	529	2	AC2112	probable homeobox
420	40.5	24.1	212	1	AB2419	phosphoglycerate m	493	40	24.1	532	2	T05281	ABC transporter, p
421	40.5	24.1	215	1	VCWGPB	coat protein - pap	494	40	24.1	535	2	P90418	ABC transporter, p
422	40.5	24.1	222	2	B84410	hypothetical prote	495	40	24.1	544	2	S58532	mark protein (trnk
423	40.5	24.1	222	2	JC1199	alpha-glucosidase	496	40	24.1	545	2	D87259	phosphoglucomutase
424	40.5	24.1	235	2	D65928	probable holoctoc	497	40	24.1	546	2	D84555	probable protein k
425	40.5	24.1	243	2	G95928	hypothetical prote	498	40	24.1	552	2	C87259	lysyl-tRNA synthet
426	40.5	24.1	245	2	F91253	probable DNA modif	499	40	24.1	554	2	AH2321	hypothetical prote
427	40.5	24.1	249	1	SYECDG	phosphatidate cyti	500	40	24.1	566	2	D84872	malose ABC transp
428	40.5	24.1	249	1	A99651	CDP-diacylglycerid	501	40	24.1	566	2	B39697	ABC transporter AT
429	40.5	24.1	249	1	AG5502	CDP-diacylglycerid	502	40	24.1	615	2	D72284	GTPase-activating
430	40.5	24.1	249	1	AG5502	hypothetical prote	503	40	24.1	633	2	AF2493	olipeptide ABC t
431	40.5	24.1	253	2	AG2362	hypothetical prote	504	40	24.1	651	2	F83376	GTPase-activating
432	40.5	24.1	263	1	E26625	oleoyl-lacetyl-carri	505	40	24.1	663	2	A39897	conserved hypothet
433	40.5	24.1	268	2	G70324	2-oxo-hepta-3-ene-	506	40	24.1	664	2	F83376	cyclomaltodextrin
434	40.5	24.1	272	2	E75607	prolipo-protein dia	507	40	24.1	718	1	ALBSMX	hypothetical prote
435	40.5	24.1	278	2	E83453	hypothetical prote	508	40	24.1	732	2	H72519	glycosyltransferas
436	40.5	24.1	282	2	A10128	phosphatidate cyti	509	40	24.1	732	2	A84107	hypothetical prote
437	40.5	24.1	285	2	AH0529	phosphatidate cyti	510	40	24.1	733	2	S76856	hypothetical prote
438	40.5	24.1	286	2	A10288	pyridoxal kinase (511	40	24.1	743	2	D84854	UvrA-like ABC tran
439	40.5	24.1	296	2	D87525	conserved hypothet	512	40	24.1	752	2	T35244	secreted protease
440	40.5	24.1	296	2	H83749	dehydrogenase/redu	513	40	24.1	781	2	G96991	NADH dehydrogenas
441	40.5	24.1	322	2	AC1929	hypothetical prote	514	40	24.1	787	2	H70374	coat protein - lei
442	40.5	24.1	329	2	D83778	hypothetical prote	515	40	24.1	791	2	B46171	hypothetical prote
443	40.5	24.1	330	2	T45981	hypothetical prote	516	40	24.1	794	2	S50687	probable dehydroge
444	40.5	24.1	336	2	G71151	probable dTPP-giuc	517	40	24.1	825	2	F95633	hypothetical prote
445	40.5	24.1	339	2	T34406	hypothetical prote	518	40	24.1	829	2	B96640	hypothetical prote
446	40.5	24.1	343	2	E69788	DNA restriction ho	519	40	24.1	840	2	T21333	pol polyprotein -
447	40.5	24.1	346	2	E68678	involved in polyke	520	40	24.1	843	2	S33123	multifunctional be
448	40.5	24.1	346	2	T08864	hypothetical prote	521	40	24.1	894	2	S54786	genome polyprotein
449	40.5	24.1	348	2	G02297	gene N33 protein -	522	40	24.1	929	2	A44048	chordin precursor
450	40.5	24.1	359	2	T16350	hypothetical prote	523	40	24.1	941	1	A53195	hypothetical WD-40
451	40.5	24.1	370	2	T35416	hypothetical prote	524	40	24.1	962	2	S58107	env polyprotein -
452	40.5	24.1	371	2	T40287	hypothetical prote	525	40	24.1	985	1	VCLJSE	env protein - simi
453	40.5	24.1	388	2	C82701	lipic acid synthe	526	40	24.1	1163	2	S18739	light junction pro
454	40.5	24.1	388	2	A82903	conserved hypothet	527	40	24.1	1163	2	JEO366	probable membrane
455	40.5	24.1	400	1	ZBBE14	44.1K zinc-binding	528	40	24.1	1246	1	S60954	major capsid prote
456	40.5	24.1	403	2	H87444	hypothetical prote	529	40	24.1	1374	1	VCBE17	protoporphyrin IX
457	40.5	24.1	403	2	T08471	harpin - Erwina a	530	40	24.1	1382	2	T01789	hypothetical prote
458	40.5	24.1	404	2	A84211	aspartate transami	531	40	24.1	1396	2	T10627	glutamate synthase
459	40.5	24.1	407	2	JC5124	hypothetical prote	532	40	24.1	1499	2	A89813	spectrin beta-6 ch
460	40.5	24.1	418	2	C72394	hypothetical prote	533	40	24.1	2364	1	A44159	utrophin - human
461	40.5	24.1	420	2	E70914	probable lipo prot	534	40	24.1	3433	1	S28381	tegument protein 2
462	40.5	24.1	423	2	C70582	probable PPE prote	535	40	24.1	3534	2	T42567	hypothetical prote
463	40.5	24.1	427	2	D81784	probable integral	536	40	24.1	3536	2	AH2515	hypothetical prote
464	40.5	24.1	428	2	G81208	AMG-related prote	537	40	23.8	4936	2	B95418	hypothetical prote
465	40.5	24.1	432	2	B69187	conserved hypothet	538	39.5	23.8	148	2	B83756	stellate protein -
466	40.5	24.1	432	2	A90465	hypothetical prote	539	39.5	23.8	172	2	S24397	hypothetical prote
467	40.5	24.1	446	2	B70776	probable gln2 - My	540	39.5	23.8	173	2	B72298	hypothetical prote

541	39.5	23.8	182	2	C83600	hypothetical prote	614	39.5	23.8	877	2	S65057	alpha-glucosidase
542	39.5	23.8	187	2	H64995	hypothetical prote	615	39.5	23.8	888	2	A55318	serine/threonine p
543	39.5	23.8	187	2	B91021	hypothetical prote	616	39.5	23.8	888	2	JC5399	dual leucine zippe
544	39.5	23.8	187	2	A85865	hypothetical prote	617	39.5	23.8	972	2	T22488	hypothetical prote
545	39.5	23.8	212	2	AD1958	hypothetical prote	618	39.5	23.8	1086	2	T05407	hypothetical prote
546	39.5	23.8	213	2	T49938	hypothetical prote	619	39.5	23.8	1095	2	G96746	hypothetical prote
547	39.5	23.8	215	2	AE0793	conserved hypotet	620	39.5	23.8	1169	2	F95000	hypothetical prote
548	39.5	23.8	229	2	S72165	hypothetical prote	621	39.5	23.8	1169	2	F95000	transcription-repa
549	39.5	23.8	231	2	E98256	hypothetical prote	622	39.5	23.8	1169	2	F95000	conserved hypotet
550	39.5	23.8	231	2	AF3028	hypothetical prote	623	39.5	23.8	1616	2	S62504	conserved hypotet
551	39.5	23.8	241	2	AH3628	nitrate reductase	624	39.5	23.8	4735	2	T17463	riboavin polyketi
552	39.5	23.8	243	2	E75517	hypothetical prote	625	39.5	23.8	7576	2	T17428	EK506 polyketide s
553	39.5	23.8	249	2	B84399	hypothetical prote	626	39.5	23.5	56	2	A82565	hypothetical prote
554	39.5	23.8	255	2	AB1942	hypothetical prote	627	39.5	23.5	103	2	F90792	IS30 transposase l
555	39.5	23.8	259	2	E84014	hypothetical prote	628	39.5	23.5	103	2	B85602	partial probable t
556	39.5	23.8	270	2	AB7537	NADPH ferredoxin r	629	39.5	23.5	128	2	S75259	hypothetical prote
557	39.5	23.8	270	2	AB7537	NADPH ferredoxin r	630	39.5	23.5	132	2	JQ2270	hypothetical 15..2K
558	39.5	23.8	306	2	AB0069	melanocortin recep	631	39.5	23.5	132	2	AG1996	hypothetical prote
559	39.5	23.8	323	2	S36636	melanocortin 3 rec	632	39.5	23.5	138	2	S74872	hypothetical prote
560	39.5	23.8	332	2	S19726	casein kinase II (633	39.5	23.5	138	2	S75128	hypothetical prote
561	39.5	23.8	332	2	S31098	casein kinase II (634	39.5	23.5	139	2	T42898	hypothetical prote
562	39.5	23.8	333	2	S31098	casein kinase II (635	39.5	23.5	143	2	E95874	conserved hypotet
563	39.5	23.8	333	2	T45853	CASEIN KINASE II,	636	39.5	23.5	143	2	JN0568	single-stranded m
564	39.5	23.8	333	2	T45853	casein kinase II (637	39.5	23.5	153	2	T12960	hypothetical prote
565	39.5	23.8	334	2	C84620	hypothetical prote	638	39.5	23.5	153	2	E95908	hypothetical prote
566	39.5	23.8	334	2	A32430	L-lactate dehydrog	639	39.5	23.5	169	2	E95908	hypothetical prote
567	39.5	23.8	371	2	S60903	hypothetical prote	640	39.5	23.5	171	2	H64602	hypothetical prote
568	39.5	23.8	378	2	A97469	hypothetical prote	641	39.5	23.5	192	2	C70006	hypothetical prote
569	39.5	23.8	385	2	AE2687	hypothetical prote	642	39.5	23.5	193	2	A33562	conserved hypotet
570	39.5	23.8	401	2	H83506	probable MFS trans	643	39.5	23.5	198	2	C84237	hypothetical prote
571	39.5	23.8	401	2	H84483	hypothetical prote	644	39.5	23.5	202	2	H97635	hypothetical prote
572	39.5	23.8	406	2	JC4600	isocitrate dehydro	645	39.5	23.5	202	2	AB2859	hypothetical prote
573	39.5	23.8	407	2	G75268	hypothetical prote	646	39.5	23.5	214	2	AC3726	DNA-binding protei
574	39.5	23.8	412	2	T26606	phosphoglycerate k	647	39.5	23.5	216	2	H66916	probable transcrip
575	39.5	23.8	422	2	C82912	hypothetical prote	648	39.5	23.5	217	1	WH02C	probable regulator
576	39.5	23.8	432	2	B84620	hypothetical prote	649	39.5	23.5	229	2	G97507	homeotic protein H
577	39.5	23.8	434	2	H89450	protein T04G9.3 (l	650	39.5	23.5	236	2	G71047	hypothetical prote
578	39.5	23.8	437	2	S53592	nmsh protein - fruit	651	39.5	23.5	237	2	JN0302	hypothetical 237 p
579	39.5	23.8	469	2	B70486	hypothetical prote	652	39.5	23.5	240	2	S73922	C89774
580	39.5	23.8	474	2	AF0759	cobalrinic acid A/C	653	39.5	23.5	246	2	C89774	C89774
581	39.5	23.8	484	2	T07782	cytochrome-c oxida	654	39.5	23.5	251	2	A91067	hypothetical prote
582	39.5	23.8	500	2	G70569	probable alpha - M	655	39.5	23.5	251	2	A85911	transposase of ins
583	39.5	23.8	504	2	T16903	hypothetical prote	656	39.5	23.5	251	2	A85911	IS30 transposase l
584	39.5	23.8	505	2	A24707	cytochrome-c oxida	657	39.5	23.5	257	2	AE2891	thiamin biosynthes
585	39.5	23.8	509	2	T11043	cytochrome-c oxida	658	39.5	23.5	261	2	AG0661	probable motility
586	39.5	23.8	510	2	T11882	cytochrome-c oxida	659	39.5	23.5	261	2	G85511	mbha protein - Bsc
587	39.5	23.8	523	2	T13137	cytochrome-c oxida	660	39.5	23.5	261	2	G64747	prolipoprotein dia
588	39.5	23.8	523	2	S62706	cytochrome-c oxida	661	39.5	23.5	266	2	E12286	acetyltransferase-
589	39.5	23.8	525	2	S26034	cytochrome-c oxida	662	39.5	23.5	285	2	S76841	hypothetical prote
590	39.5	23.8	525	2	S26022	cytochrome-c oxida	663	39.5	23.5	294	2	A44856	haloacetate dehalo
591	39.5	23.8	526	2	S62762	cytochrome-c oxida	664	39.5	23.5	294	2	A44856	hypothetical prote
592	39.5	23.8	528	2	E71525	probable ABC trans	665	39.5	23.5	296	2	T47062	hypothetical prote
593	39.5	23.8	530	2	E72129	ABC transporter pr	666	39.5	23.5	296	2	AC0231	probable ABC trans
594	39.5	23.8	530	2	A86494	ABC transporter pr	667	39.5	23.5	299	2	S62400	gene O protein - p
595	39.5	23.8	533	2	AE5931	cytochrome-c oxida	668	39.5	23.5	304	2	E30777	replication protei
596	39.5	23.8	533	2	AE1616	ABC transporter (A	669	39.5	23.5	309	2	E30777	heat-shock protei
597	39.5	23.8	536	2	AG1253	ABC transporter (A	670	39.5	23.5	311	2	D72759	hypothetical precu
598	39.5	23.8	536	2	T07946	cytochrome-c oxida	671	39.5	23.5	311	2	S09575	osteopontin precu
599	39.5	23.8	541	2	A46327	cytochrome-c oxida	672	39.5	23.5	314	1	A54522	probable ABC trans
600	39.5	23.8	546	2	H72357	hypothetical prote	673	39.5	23.5	314	2	A54522	hypothetical prote
601	39.5	23.8	577	2	B75585	probable long-chain	674	39.5	23.5	318	2	D69742	alkaline proteinas
602	39.5	23.8	587	2	C86744	myosin-crossreacti	675	39.5	23.5	321	1	S27501	hypothetical prote
603	39.5	23.8	589	2	AC3370	hypothetical prote	676	39.5	23.5	325	2	T05512	transcription-repa
604	39.5	23.8	605	2	H87468	ubiquinol oxidase	677	39.5	23.5	326	2	AD3434	conserved hypotet
605	39.5	23.8	605	2	JC2363	protein kinase (EC	678	39.5	23.5	337	2	A45038	riboavin polyketi
606	39.5	23.8	709	2	E64213	DNA topoisomerase	679	39.5	23.5	360	2	T43404	casein kinase II (
607	39.5	23.8	723	1	JN0531	p-aminobenzoic aci	680	39.5	23.5	375	2	S58784	probable membrane
608	39.5	23.8	751	1	WMKRGB	probable core prot	681	39.5	23.5	377	2	AH3372	muconate cyclisom
609	39.5	23.8	793	1	E64545	hypothetical prote	682	39.5	23.5	383	1	F65241	transposase - Esch
610	39.5	23.8	811	2	S17908	hypothetical prote	683	39.5	23.5	384	1	T22434	phosphoprotein pho
611	39.5	23.8	820	2	A86247	hypothetical prote	684	39.5	23.5	385	2	A45827	N-acetylornithine
612	39.5	23.8	840	1	A43970	H+-exporting ATPas	685	39.5	23.5	387	2	E96701	pigment production
613	39.5	23.8	873	1	S53828	cytochrome-c oxida	686	39.5	23.5	390	2	E96701	unknown protein, 8

687	39	23.5	394	2	710200	hypothetical prote	760	39	23.5	686	2	F87489	NADH dehydrogenase
688	39	23.5	395	2	T05680	hypothetical prote	761	39	23.5	698	2	F64839	yeast protein precu
689	39	23.5	396	2	A85808	hypothetical prote	762	39	23.5	698	2	D90771	hypothetical prote
690	39	23.5	408	2	T31254	plasmalogen activa	763	39	23.5	698	2	H83633	hypothetical prote
691	39	23.5	415	2	S20047	chlorohydrolyase [i	764	39	23.5	702	2	A81619	uncharacterized co
692	39	23.5	428	2	E87573	hypothetical prote	765	39	23.5	709	2	A97218	tomato-dependent rec
693	39	23.5	435	2	T15290	hypothetical prote	766	39	23.5	712	2	B87683	cleavage stimulat
694	39	23.5	440	2	S40410	cholesterol synthas	767	39	23.5	717	2	S50852	methyl-accepting c
695	39	23.5	442	2	S68738	nitrate-binding pr	768	39	23.5	735	2	B87601	probable ATP-bind
696	39	23.5	444	2	D64636	proteinase (EC 3.4	769	39	23.5	740	1	T02567	cryptophan 2-mono
697	39	23.5	445	2	AF0483	metalloproteinase	770	39	23.5	755	1	DAA6WT	conserved hypotet
698	39	23.5	445	2	T49943	GDP dissociation 1	771	39	23.5	761	2	G70393	repair/recombinati
699	39	23.5	446	2	G72731	probable malate ox	772	39	23.5	854	2	S61164	hypothetical prote
700	39	23.5	446	2	S77389	nitrate transport	773	39	23.5	855	2	G86189	hypothetical prote
701	39	23.5	448	2	D71477	glutamate transpor	774	39	23.5	913	3	JG0168	gob-5 protein - mo
702	39	23.5	449	2	S71005	ornithine decarbox	775	39	23.5	971	2	H97454	hypothetical prote
703	39	23.5	450	1	DCCHO	acetyl-CoA carboxy	776	39	23.5	971	2	AB2673	conserved hypotet
704	39	23.5	454	2	AE1244	ornithine decarbox	777	39	23.5	994	2	A70776	probable gln - My
705	39	23.5	455	1	DCMYOC	malate oxidoreduct	778	39	23.5	1004	2	H87712	glutamate-ammonia
706	39	23.5	456	2	F69465	aspartate transami	779	39	23.5	1075	2	T47603	beta Galactosidase
707	39	23.5	456	2	T06136	biphenyl dioxyena	780	39	23.5	1161	2	S18738	pol protein - siml
708	39	23.5	458	1	A42409	ig mu chain C regi	781	39	23.5	1165	2	A70423	valine-tRNA ligase
709	39	23.5	458	1	MHRM	ig mu chain C regi	782	39	23.5	1165	2	D59433	C. elegans protein
710	39	23.5	467	2	G83843	thiaminokinase BHI	783	39	23.5	1184	2	G71262	probable pyruvate
711	39	23.5	471	2	AH1104	cysteinyl-tRNA syn	784	39	23.5	1270	2	T22615	hypothetical prote
712	39	23.5	471	2	AH1466	probable membrane	785	39	23.5	1286	1	RUBOP	interphotoreceptor
713	39	23.5	473	2	AG0948	probable UDP-N-ace	786	39	23.5	1289	2	D87789	protein C34G6.4 [i
714	39	23.5	475	2	C81351	Penicillin-binding	787	39	23.5	1289	2	F772308	hypothetical prote
715	39	23.5	477	2	AC0903	probable protein k	788	39	23.5	1409	2	T29594	hypothetical prote
716	39	23.5	477	2	C84871	ig mu chain C regi	789	39	23.5	1581	1	VGMJBY	peplomeric glycoprot
717	39	23.5	479	1	MHRM	ig mu chain C regi	790	39	23.5	1622	2	D86428	glutathione S-conj
718	39	23.5	482	2	AC0426	serine-type D-Ala-	791	39	23.5	1704	2	S71363	probable ATP-bind
719	39	23.5	482	2	A48370	nitrogen fixation	792	39	23.5	1704	2	A59188	ATP-binding cassel
720	39	23.5	495	2	C63122	probable aldehyde	793	39	23.5	1804	2	S56247	probable membrane
721	39	23.5	496	2	E90936	probable transport	794	39	23.5	1822	2	T50207	probable protein t
722	39	23.5	496	2	A85785	probable membrane	795	39	23.5	1932	2	AF1946	hypothetical hydr
723	39	23.5	496	2	AF0165	hypothetical prote	796	39	23.5	2034	2	T22147	hypothetical prote
724	39	23.5	496	2	C64935	probable aldehyde	797	39	23.5	2228	2	E97542	beta-galactosidase
725	39	23.5	498	2	F95358	prolyl-tRNA synthet	798	39	23.5	2233	2	B95075	hypothetical prote
726	39	23.5	499	2	D75416	protein F25G6.7 [i	799	39	23.5	2326	2	T29140	probable sensor-11
727	39	23.5	501	2	B89135	tidl homolog Rv231	800	39	23.5	2344	2	T41590	enoyl-lacyl-carrie
728	39	23.5	505	1	D70703	aldehyde dehydroge	801	39	23.5	2512	1	XYCHFA	polypeptide - dour
729	39	23.5	507	2	B87400	hypothetical prote	802	39	23.5	3005	2	T08841	plasma membrane ca
730	39	23.5	507	2	T34152	hypothetical prote	803	39	23.5	57	2	B49570	Ca2+-transporting
731	39	23.5	516	2	T00791	purple acid phosph	804	39	23.2	130	2	A44301	conserved hypotet
732	39	23.5	523	2	A97576	3297 (imported) -	805	39	23.2	189	2	F75358	potassium-transpor
733	39	23.5	523	2	A12796	hypothetical prote	806	39	23.2	187	2	H87446	hypothetical prote
734	39	23.5	529	2	B81186	L-lactate permease	807	39	23.2	205	2	C72531	hypothetical prote
735	39	23.5	529	2	F69020	conserved hypotet	808	39	23.2	221	2	T11797	nitrate reductase
736	39	23.5	532	1	A40876	dimethylalliline mo	809	39	23.2	225	1	RDECNG	nitrate reductase
737	39	23.5	532	1	A33768	flavin-containing	810	39	23.2	225	2	D90845	nitrate reductase
738	39	23.5	532	2	S33758	hypothetical prote	811	39	23.2	225	2	C85703	probable membrane
739	39	23.5	541	2	T11554	phosphoglucumutase	812	39	23.2	225	2	S52501	hypothetical prote
740	39	23.5	547	2	E70650	hypothetical prote	813	39	23.2	233	2	S76857	hypothetical prote
741	39	23.5	554	2	T34028	probable acetolact	814	39	23.2	234	2	T11750	hypothetical prote
742	39	23.5	556	2	E72483	protein K02R2.4 [i	815	39	23.2	243	2	S38236	pyruvate dehydroge
743	39	23.5	556	2	D88700	hypothetical prote	816	39	23.2	254	2	AF0415	conserved hypotet
744	39	23.5	558	2	T47851	protein MD1G5.5 [i	817	39	23.2	259	2	F82707	probable bis(5'-nu
745	39	23.5	559	2	F68392	hypothetical prote	818	39	23.2	269	2	H82322	myelin proteolipid
746	39	23.5	559	2	T33425	amino-acid permea	819	39	23.2	277	2	F51370	hypothetical prote
747	39	23.5	567	2	T50051	cytochrome P450 en	820	39	23.2	285	2	T16139	probable thiosulfa
748	39	23.5	574	2	JC7327	penicillin-binding	821	39	23.2	285	2	A87000	conserved hypotet
749	39	23.5	575	2	G87681	ABC transporter sl	822	39	23.2	322	2	A83549	hypothetical prote
750	39	23.5	585	2	S77114	hypothetical prote	823	39	23.2	342	2	T29245	anthranilate phosph
751	39	23.5	590	2	H71977	hypothetical prote	824	39	23.2	362	2	AF1950	nitrogen regulatio
752	39	23.5	593	2	H64528	hypothetical prote	825	39	23.2	380	2	AC3360	aspartate aminotra
753	39	23.5	623	2	H83399	guanine-protein alcoh	826	39	23.2	380	2	H84320	opsin 2 - fruit fl
754	39	23.5	627	2	T47968	dynamn-like prote	827	39	23.2	381	1	COEF2	hypothetical prote
755	39	23.5	647	2	I64022	hypothetical prote	828	39	23.2	382	2	T27076	bicyclic mycin resis
756	39	23.5	661	2	A69252	3-hydroxyacyl-CoA	829	39	23.2	402	2	E69783	MFS permease (limp
757	39	23.5	664	2	B36885	tau-type ubiquitin	830	39	23.2	411	2	AB2903	tryptophan synthas
758	39	23.5	673	2	F87636	TPP domain protein	831	39	23.2	422	2	H72363	conserved hypotet
759	39	23.5	686	1	ALDYAT	amylase A (EC 3.2.	832	39	23.2	427	2	H75638	

833	38.5	23.2	437	2	E47070	906	38	22.9	190	2	AH2216	nitrate transport
834	38.5	23.2	437	2	AB0378	907	38	22.9	201	2	HA1131	conserved hypotnet
835	38.5	23.2	438	2	A72430	908	38	22.9	201	2	B81692	hypothetical prote
836	38.5	23.2	454	2	AF3537	909	38	22.9	204	2	T05700	hypothetical prote
837	38.5	23.2	465	2	G84567	910	38	22.9	215	2	S56540	flm1 protein - Esc
838	38.5	23.2	466	2	F95041	911	38	22.9	215	2	E86129	hypothetical prote
839	38.5	23.2	471	1	T04935	912	38	22.9	215	2	B91288	fibrial protein [
840	38.5	23.2	472	1	S55379	913	38	22.9	217	2	F69979	caffeoyl-CoA O-met
841	38.5	23.2	478	2	C97678	914	38	22.9	219	2	T01186	protein kinase (EC
842	38.5	23.2	484	2	JC5282	915	38	22.9	229	2	A49101	enolase-phosphatas
843	38.5	23.2	485	2	S74708	916	38	22.9	231	2	S76372	hypothetical prote
844	38.5	23.2	488	2	T20124	917	38	22.9	236	2	JC5664	guanidinocetate N
845	38.5	23.2	489	1	S46671	918	38	22.9	238	2	B83799	required for a lat
846	38.5	23.2	497	2	C91206	919	38	22.9	239	2	T45650	hypothetical prote
847	38.5	23.2	497	2	E86052	920	38	22.9	241	2	A71020	hypothetical prote
848	38.5	23.2	497	2	G65169	921	38	22.9	244	2	A75408	tiaseophosphate is
849	38.5	23.2	509	2	A36392	922	38	22.9	245	2	A83409	hypothetical prote
850	38.5	23.2	510	2	E84347	923	38	22.9	246	2	S50731	hypothetical prote
851	38.5	23.2	513	2	T43334	924	38	22.9	248	2	E83823	hypothetical prote
852	38.5	23.2	517	2	S01501	925	38	22.9	250	2	D84820	ABC transporter (A
853	38.5	23.2	519	2	T11129	926	38	22.9	257	2	T47689	probable expansin
854	38.5	23.2	519	2	D70695	927	38	22.9	260	2	AE3367	3-demethylubiquino
855	38.5	23.2	547	2	AF0047	928	38	22.9	276	2	AE3367	hypothetical prote
856	38.5	23.2	560	2	A86214	929	38	22.9	277	2	T24329	conserved hypotnet
857	38.5	23.2	566	2	UH0218	930	38	22.9	279	2	G97143	probable membrane
858	38.5	23.2	576	2	E86208	931	38	22.9	289	2	AH2088	heat shock protein
859	38.5	23.2	604	2	F64081	932	38	22.9	297	2	S17474	hypothetical prote
860	38.5	23.2	605	2	E71253	933	38	22.9	309	2	T24510	hypothetical prote
861	38.5	23.2	620	2	T15273	934	38	22.9	310	2	F64146	hypothetical prote
862	38.5	23.2	621	2	G87459	935	38	22.9	312	2	AD0779	hypothetical prote
863	38.5	23.2	671	2	AF1294	936	38	22.9	317	2	AD0779	conserved hypotnet
864	38.5	23.2	673	2	H72563	937	38	22.9	317	2	D81181	probable uridine p
865	38.5	23.2	790	2	G90477	938	38	22.9	322	2	T20018	hypothetical prote
866	38.5	23.2	815	2	H82137	939	38	22.9	323	2	T44256	abc transporter, p
867	38.5	23.2	818	2	C72671	940	38	22.9	323	2	AE3630	thiamin biosynthes
868	38.5	23.2	838	2	S49750	941	38	22.9	326	2	AB3198	noxi1 protein [imp
869	38.5	23.2	846	2	T19179	942	38	22.9	327	1	S76143	hypothetical prote
870	38.5	23.2	859	2	S70584	943	38	22.9	328	1	DMHUT	probable aldehyde
871	38.5	23.2	868	1	UC4283	944	38	22.9	330	2	B37756	L-serine ammonia-1
872	38.5	23.2	916	2	F71962	945	38	22.9	331	2	A84389	liru protein - Esc
873	38.5	23.2	919	2	T16693	946	38	22.9	331	2	T15458	ribonucleoside red
874	38.5	23.2	924	2	C83572	947	38	22.9	331	2	T44514	hypothetical prote
875	38.5	23.2	937	2	PO0864	948	38	22.9	333	2	D85068	hypothetical prote
876	38.5	23.2	1009	2	AF2033	949	38	22.9	334	2	E83122	D133-like protein
877	38.5	23.2	1047	2	T41343	950	38	22.9	340	2	A10195	probable iron/asco
878	38.5	23.2	1151	2	E86826	951	38	22.9	342	2	A45966	DNA-directed DNA p
879	38.5	23.2	1474	2	F69009	952	38	22.9	354	2	T45837	Ig alpha chain C r
880	38.5	23.2	1577	2	A35140	953	38	22.9	357	2	I55210	hypothetical prote
881	38.5	23.2	2219	2	T27684	954	38	22.9	358	2	E95842	tricarboxylate car
882	38.5	23.2	3224	1	S58884	955	38	22.9	360	2	E95126	probable lactose t
883	38.5	23.2	3712	1	YGCEVC	956	38	22.9	362	2	AC1906	aminotransferase,
884	38	22.9	40	2	T08870	957	38	22.9	363	2	S34734	Chorismate synthas
885	38	22.9	86	2	H97052	958	38	22.9	364	2	E83276	ATP-binding protei
886	38	22.9	90	2	D82760	959	38	22.9	370	2	C81780	hypothetical prote
887	38	22.9	92	2	H82371	960	38	22.9	371	2	AE7997	quinolinate synthe
888	38	22.9	122	2	G84579	961	38	22.9	373	1	KMECTD	pyridoxal-phosphat
889	38	22.9	135	2	B95102	962	38	22.9	373	1	S29934	chorismate mutase
890	38	22.9	135	2	D97970	963	38	22.9	373	2	C85906	chorismate mutase-
891	38	22.9	138	2	T49392	964	38	22.9	373	2	AB0399	prephenate dehydro
892	38	22.9	141	2	T34511	965	38	22.9	373	2	A10832	hypothetical prote
893	38	22.9	159	2	C73446	966	38	22.9	373	2	A86859	hypothetical prote
894	38	22.9	169	2	C69394	967	38	22.9	375	2	AH0209	probable sugar ABC
895	38	22.9	172	2	I48171	968	38	22.9	376	2	S63613	probable ATP-bindi
896	38	22.9	173	2	S69468	969	38	22.9	377	2	G86724	acetylornithine tr
897	38	22.9	175	1	CYBOAB	970	38	22.9	377	2	F90523	conserved hypotnet
898	38	22.9	175	1	CYH7AB	971	38	22.9	380	1	VCYGR	coat protein - Cym
899	38	22.9	175	2	A53871	972	38	22.9	380	1	S05457	coat protein - Cym
900	38	22.9	175	2	A39608	973	38	22.9	381	2	AG3148	hypothetical prote
901	38	22.9	175	2	I53319	974	38	22.9	388	2	AF1541	antibiotic resista
902	38	22.9	175	2	A23681	975	38	22.9	388	2	AH1183	antibiotic resista
903	38	22.9	175	2	JC5971	976	38	22.9	388	2	T29364	hypothetical prote
904	38	22.9	175	2	H97829	977	38	22.9	392	2	C84960	tRNA (5-methylam
905	38	22.9	183	2	A41302	978	38	22.9				

979	38	22.9	392	2	B66549	polymorphic outer
980	38	22.9	392	2	B12075	polymorphic outer
981	38	22.9	394	2	T45672	hypothetical prote
982	38	22.9	395	2	A12412	hypothetical prote
983	38	22.9	396	2	T50229	probable transmem
984	38	22.9	397	2	S28274	hypothetical prote
985	38	22.9	398	2	D66700	hypothetical prote
986	38	22.9	399	2	F70937	hypothetical prote
987	38	22.9	400	2	A26258	endoplasmic retic
988	38	22.9	401	1	S10528	5-aminolevulinat
989	38	22.9	401	1	T21288	phosphoprotein pho
990	38	22.9	401	2	A50696	probable integral
991	38	22.9	402	2	T24710	probable transpos
992	38	22.9	402	2	A50193	transposase, IS285
993	38	22.9	402	2	A50356	transposase, IS285
994	38	22.9	402	2	A50472	transposase, IS285
995	38	22.9	402	2	A50341	transposase, IS285
996	38	22.9	402	2	AD0002	transposase, IS285
997	38	22.9	402	2	A50190	transposase, IS285
998	38	22.9	402	2	A50242	transposase, IS285
999	38	22.9	402	2	A50267	transposase, IS285
1000	38	22.9	402	2	A50348	transposase, IS285

ALIGNMENTS

RESULT 1

UC7350 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
C:Species: Mus musculus (house mouse)
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: J07350
R:Uchimura, K.; Fasaki, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa, N.;
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
A:Reference number: UC7350; MUID:20374462; PMID:10913333
A:Accession: J07350
A:Molecule type: mRNA
A:Residues: 1-484 <UCH>
A:Cross-references: DDBJ:AB040710
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 80.7%; Score 134; DB 2; Length 484;
Best Local Similarity 72.4%; Pred. No. 3.2e-11;
Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAHV 29

DB 107 TWRGSSFLGELFNQHPDVFYLYEPWHL 135

RESULT 2

J07351 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
C:Species: Homo sapiens (man)
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: J07351
R:Uchimura, K.; Fasaki, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa, N.;
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
A:Reference number: J07350; MUID:20374462; PMID:10913333
A:Accession: J07351
A:Molecule type: mRNA
A:Residues: 1-486 <UCH>
A:Cross-references: DDBJ:AB040711
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 80.7%; Score 134; DB 2; Length 486;

Best Local Similarity 72.4%; Pred. No. 3.2e-11;
Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAHV 29

DB 109 TWRGSSFLGELFNQHPDVFYLYEPWHL 137

RESULT 3

JE0261 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
C:Species: Homo sapiens (man)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
C:Accession: JE0261
R:Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuo
T.;
J. Biochem. 124, 670-678, 1998
A:Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis
A:Reference number: JE0261; MUID:98391845; PMID:9722682
A:Accession: JE0261
A:Molecule type: mRNA
A:Residues: 1-484 <UCH>
A:Cross-references: DDBJ:AB014679
C:Comment: This protein catalyzes the transfer of sulfur of sulfate from 3'-phosphoad
C:Superfamily: chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match 74.7%; Score 124; DB 2; Length 484;
Best Local Similarity 69.0%; Pred. No. 8.6e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAHV 29

DB 126 TWRGSSFLGELFNQHPDVFYLYEPWHL 154

RESULT 4

A57397 chondroitin 6-sulfotransferase (EC 2.8.2.17) precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 20-Jun-2000
C:Accession: A57397
R:Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuc
J. Biol. Chem. 270, 18575-18580, 1995
A:Title: Molecular cloning and expression of chick chondrocyte chondroitin 6-sulfot
A:Reference number: A57397; MUID:95355450; PMID:7629189
A:Accession: A57397
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-458 <EUC>
A:Cross-references: GB:D49915; NID:9971262; PIDN:BA08655.1; PID:9971263
C:Superfamily: chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match 60.8%; Score 101; DB 2; Length 458;
Best Local Similarity 59.3%; Pred. No. 1.5e-06;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFFVGOLFQHPDVFYIMEPAHV 29

DB 122 RTGSSFLGELFNQHPDVFYLYEPWHL 148

RESULT 5

AC2012 hypothetical protein al11649 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AC2012
R:Kanevo, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

AE2599
 hypothetical protein Atu0187 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
 C:Accession: AE2599
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AE2599
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-614 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AA141211.1; PID:g17738513; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0187
 A:Map position: circular chromosome
 C:Superfamily: periplasmic oligopeptide-binding protein

Query Match 31.3%; Score 52; DB 2; Length 614;
 Best Local Similarity 57.9%; Pred. No. 20;
 Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

OY 2 WRSSSFVQGLFGQHPDV 20
 |||||
 Db 33 WRKGISTVQEL--KHPDGF 49

RESULT 11
 AG2164
 hypothetical protein all2870 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AG2164
 R:Kaneko, T.; Nakamura, Y.; Molk, G.P.; Kurita, T.; Sasaoka, S.; Watanabe, A.; Itiguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AG2164
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-641 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA074569.1; PID:g17131964; GSPDB:GN00179
 C:Genetics:
 A:Gene: all2870

Query Match 31.3%; Score 52; DB 2; Length 641;
 Best Local Similarity 46.7%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 WRSSSFVQGLFGQHPDV 16
 |||||
 Db 298 WRNGESYINSFQYH 312

RESULT 12
 T48933
 WD repeat domain protein - Arabidopsis thaliana
 N:Alternate names: protein F14L2.80
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T48933
 R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z25008

A:Accession: T48933
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1051 <JOR>
 A:Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.80
 A:Experimental source: cultivar Columbia; BAC clone F14L2
 C:Genetics:
 A:Gene: ATSP:F14L2.80
 A:Map position: 3
 A:Introns: 36/3; 121/3; 205/3; 336/2; 668/3; 739/1; 773/1; 853/2; 894/2; 932/3

Query Match 30.1%; Score 50; DB 2; Length 1051;
 Best Local Similarity 36.7%; Pred. No. 72;
 Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

OY 1 WRSSSFVQGLFGQHPDV--FYLMEPAWH 28
 |||||
 Db 823 SWNLGAPGELAGLVDRYKMARKPWN 852

RESULT 13
 A37792
 Spectrin beta-H chain - fruit fly (Drosophila melanogaster) (fragment)
 C:Species: Drosophila melanogaster
 C:Date: 30-Apr-1991 #sequence_revision 08-Nov-1996 #text_change 16-Jul-1999
 C:Accession: A37792; S70848; S15666
 R:Dubreuil, R.R.; Byers, T.J.; Stewart, C.T.; Kiehart, D.P.
 J. Cell Biol. 111, 1849-1858, 1990
 A:Title: A beta-spectrin isoform from Drosophila (beta-H) is similar in size to verte
 A:Reference number: A37792; MUID:91055599; PMID:2229176
 A:Accession: A37792
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1645 <DUB>
 A:Cross-references: GB:X53992
 A:Note: Met-14 is the probable initiator
 R:Dubreuil, R.R.
 submitted to the EMBL Data Library, July 1990
 A:Reference number: S70848
 A:Accession: S70848
 A:Molecule type: mRNA
 A:Residues: 1-400,403-410, 'OL',411-1645 <DUB>
 A:Cross-references: EMBL:X53992; MUID:97654; PID:g7655
 A:Note: Met-14 is the probable initiator
 C:Genetics:
 A:Gene: FlyBase:kst
 A:Cross-references: FlyBase:FBgn0004167
 C:Superfamily: alpha-actinin actin-binding domain homology; SH3 homology; spectrin/dy
 C:Keywords: actin binding
 F:36-258/Domain: alpha-actinin actin-binding domain homology <ACT>
 F:248-399/Domain: spectrin/dystrophin repeat homology <SP1>
 F:400-510/Domain: spectrin/dystrophin repeat homology <SP2>
 F:851-896/Domain: SH3 homology <SH3>
 F:990-1091/Domain: spectrin/dystrophin repeat homology <SP3>
 F:1303-1408/Domain: spectrin/dystrophin repeat homology <SP4>
 F:1409-1512/Domain: spectrin/dystrophin repeat homology <SP5>

Query Match 30.1%; Score 50; DB 2; Length 1645;
 Best Local Similarity 45.0%; Pred. No. 1,2e+02;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 1 WRSSSFVQGLFGQHPDV 20
 |||||
 Db 186 WRSGLGFNLIHSHRPDLF 205

RESULT 14
 C83330
 potassium uptake protein KUP PA0917 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: C83330
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

OY 8 FVQGLFGHPDVFYLMKP 25
 11 : : : : :
 DB 143 F1RALAQHAEIFFLDEP 160

RESULT 19

A:Accession: A11677
 A:Title: metal cations ABC transporter, ATP-binding protein homolog 11n1963 [imported] - *Listeria*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
 C:Accession: A11677
 R:Diasser, P.; Franjeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Meek, C.; Schlueter, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11677
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <GLA>
 A:Cross-references: GB:AF592022; PIDN:CA097193.1; PID:q16414464; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: 11n1963
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match

Best Local Similarity 28.9%; Score 48; DB 2; Length 240;
 Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 8 FVQGLFGHPDVFYLMKP 25
 11 : : : : :
 DB 143 F1RALAQHAEIFFLDEP 160

RESULT 20

D70670
 Hypothetical protein RV2959c - *Mycobacterium tuberculosis* (strain H37RV)
 N:Alternate names: U00021b protein
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: D70670; S73062
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Kellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: D70670
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-245 <COL>
 A:Cross-references: GB:Z83018; GB:AL123456; NID:93261671; PIDN:CA05417.1; PID:e283380;
 A:Experimental source: strain H37RV
 R:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, September 1994
 A:Description: *Mycobacterium tuberculosis* cosmid tb02.
 A:Reference number: S73053
 A:Accession: S73062
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-245 <SMI>
 A:Cross-references: EMBL:U00024; NID:9560506; PIDN:AAA50941.1; PID:9560520
 C:Genetics:
 A:Gene: RV2959c

Query Match 28.9%; Score 48; DB 2; Length 245;
 Best Local Similarity 69.2%; Pred. No. 27;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 WRSQSFYQGLNG 14
 11 : : : : :
 DB 5 WRSRTSLVQGLIG 17

RESULT 21

S73902
 UDP-glucose 4-epimerase galP - *Mycoplasma pneumoniae* (strain ATCC 29342)
 N:Alternate names: hypothetical protein A65_0r1338
 C:Species: *Mycoplasma pneumoniae*
 A:Valley: ATCC 29342
 C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S73902
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*
 A:Reference number: S73327; MUID:97105885; PMID:8948633
 A:Accession: S73902
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-338 <HIM>
 A:Cross-references: EMBL:AE000056; GB:U00089; NID:91674263; PIDN:AA096224.1; PID:9167
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Gene: galP
 A:Genetic code: SGC3
 C:Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo f:8-338/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 28.9%; Score 48; DB 2; Length 338;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 5 GSSEFYQGLFGHPDV 19
 11 : : : : :
 DB 18 GSCFIDQLKQYDPV 32

RESULT 22

T10801
 GDP dissociation inhibitor GDII - *Volvox carter* f. *nagariensis*
 C:Species: *Volvox carter* f. *nagariensis*
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T10801
 R:Beysen, K.; Fabry, S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Identification and characterization of a lower plant Ypt/Rab guanosine
 A:Reference number: Z17156
 A:Accession: T10801
 A:Status: preliminary; translated from GS/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-443 <BEY>
 A:Cross-references: EMBL:U62866; NID:91572517; PID:91572518
 A:Experimental source: strain HK10
 C:Genetics:
 A:Gene: GDII
 C:Function:
 A:Description: inhibits dissociation of GDP from GMP binding proteins
 C:Superfamily: human GDP dissociation inhibitor XAP-4

Query Match 28.9%; Score 48; DB 2; Length 443;
 Best Local Similarity 47.4%; Pred. No. 53;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 8 FVQGLFGHPDVFYLMKP 26
 11 : : : : :
 DB 183 F1GHALGLHDDAYLTQPA 201

RESULT 23
 E83617
 Probable aminotransferase PA0221 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E83617
 R:Stover, C.K.; Pham, X.O.; Ewld, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: E83617
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-460 <STO>
 A:Cross-references: GB:AE004460; GB:AE004091; NID:g9946055; PIDN:AA03610.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0221
 C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 28.9%; Score 48; DB 2; Length 460;
 Best Local Similarity 33.3%; Pred. No. 55;
 Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 5 GSSFGQLFGQHPDVFYLMPEAWH 28
 DB 169 GAKFMHMGMLPFAHIDEPYMY 192

RESULT 24
 F70322
 gltamyl-tRNA (Gln) amidotransferase subunit A - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 22-Jun-1999
 C:Accession: F70322
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: F70322
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-478 <AOE>
 A:Cross-references: GB:AE000680; NID:g2982948; PIDN:AA06569.1; PID:g2982954; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: gntA
 C:Superfamily: indoleacetamide hydrolase

Query Match 28.9%; Score 48; DB 2; Length 478;
 Best Local Similarity 36.7%; Pred. No. 58;
 Matches 11; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 1 SMRSGSFVGLFGQHPDVFYLMPEAWH 24
 DB 437 AKMDGLPVGGQLGKRWDETTLLQISTLWE 466

RESULT 25
 T23120
 hypothetical protein H25K10.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 R:Morimoto, B.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19684
 A:Accession: T23120
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-518 <MIL>
 A:Cross-references: EMBL:Z92796; PIDN:CAB07233.1; GSPDB:GN00022; CESP:H25K10.6
 A:Experimental source: clone H25K10

C:Genetics:
 A:Gene: CESP:H25K10.6
 A:Map position: 4
 A:Introns: 59/1; 144/1; 211/1; 307/3; 453/3
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein T05A6.4

Query Match 28.9%; Score 48; DB 2; Length 518;
 Best Local Similarity 26.9%; Pred. No. 63;
 Matches 7; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 2 WRSFGVGLFGQHPDVFYLMPEAWH 27
 DB 484 YTTNAFIDYDYGEPYTFELCELM 509

RESULT 26
 D71307
 conserved hypothetical integral membrane protein TP0582 - *Syphilis spirochete*
 C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: D71307
 R:Freaser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
 rison, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: D71307
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-495 <COI>
 A:Cross-references: GB:AE001233; GB:AE000520; NID:g3322870; PIDN:AA06556.1; PID:g332
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0582

Query Match 28.6%; Score 47.5; DB 2; Length 495;
 Best Local Similarity 35.3%; Pred. No. 70;
 Matches 12; Conservative 7; Mismatches 10; Indels 5; Gaps 2;

OY 1 SMRSGSFVGLFGQHPDVFYLMPEAWH 29
 DB 416 SMRTAIFAGQLFRTSAHVPTPIHLDPAHYL 449

RESULT 27
 G72781
 hypothetical protein APE0240 - *Aeropyrum pernix* (strain K1)
 C:Species: *Aeropyrum pernix*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: G72781
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aero*
 A:Reference number: A72450; MUID:9910339; PMID:10382966
 A:Accession: G72781
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-176 <KAN>
 A:Cross-references: DBJ:AP000058; NID:g5103388; PIDN:BA079153.1; PID:g5103632
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0240
 C:Superfamily: *Aeropyrum pernix* hypothetical protein APE0240

Query Match 28.3%; Score 47; DB 2; Length 176;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 5 GSSFGVGLFGQHPDVFYLMPEAWH 24
 DB 45 GSVFVSAVFSQHPDVFYLMPEAWH 64

A.Variety: PCC 6803
C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C.Accession: S76464
R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A.Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A.Reference number: S74322; MUID:97061201; PMID:8905231
A.Accession: S76464
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-465 <KAN>
A.Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BA18593.1; PID:d101932
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 465;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 FVGOLFQGHDPVFLM 23
DB 25 WVGQIFSQLADKFLY 40

RESULT 38
S51613
C.Species: *Candida albicans*
C.Date: 15-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 08-Sep-1995
C.Accession: S51613; S49207
R.Sherlock, G.; Bahman, A.M.; Mahal, A.; Shieh, J.C.; Ferreira, M.; Rosamond, J.
Mol. Gen. Genet. 245, 716-723, 1994
A.Title: Molecular cloning and analysis of CDC28 and cyclin homologues from the human fu
A.Reference number: S51611; MUID:95131949; PMID:7830719
A.Accession: S51613
A.Molecule type: DNA
A.Residues: 1-465 <SHE>
A.Cross-references: EMBL:X80033

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 465;
Matches 6; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 6 SSFVGOLFQGHDPVFLM 21
DB 211 SWIGELFOFIPNITY 226

RESULT 39
C83052
C.Species: *Pseudomonas aeruginosa*
C.Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C.Accession: C83052
R.Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Ba
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A.Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pat
A.Reference number: A82950; MUID:20437337; PMID:10984043
A.Accession: C83052
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-701 <STO>
A.Cross-references: GB:AE004886; GB:AE004091; NID:g9951001; PIDN:AG08126.1; GSPDB:GN001
A.Experimental source: strain PA01
C.Genetics:
A.Gene: PNP; PA4740
C:Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 701;
Matches 32.1%; Pred. No. 1.2e+02;

Matches 9; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

QY 4 SGSSFVGL-----FGHPDVFYLMPEPM 27
DB 140 SGIFPAGPIGAARVGFPEIGYILNPTY 167

RESULT 40
D96521
C.Species: *Arabidopsis thaliana* (mouse-ear cress)
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C.Accession: D96521
R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
sen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Malt, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A.Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A.Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A.Reference number: A86141; MUID:21016719; PMID:11130712
A.Accession: D96521
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1340 <STO>
A.Cross-references: GB:AE005173; NID:g8778513; PIDN:AAF9521.1; GSPDB:GN00141
C.Genetics:
A.Gene: F21D18.16
A.Map position: 1

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 1340;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 VGLFQGHDPVFLMPEPA 26
DB 658 VGDLHGQLHDLVYIMODA 675

RESULT 41
AB0480
C.Species: *Yersinia pestis*
C.Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C.Accession: AB0480
R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barr
Nature 413, 523-527, 2001
A.Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A.Reference number: AB0001; MUID:21470413; PMID:11586360
A.Accession: AB0480
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-3013 <KUR>
A.Cross-references: GB:AL590842; PIDN:CAC93406.1; PID:g15981852; GSPDB:GN00175
C.Genetics:
A.Gene: YPO3944

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 3013;
Matches 9; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 SMRSSEFVGLFGHPDVFYLMPEPM 27
DB 2942 TWMGKRAVGNLMSWGDPSAYTAPGM 2968

RESULT 42
ESDKM

oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) - mallard
N:Alternate names: S-acyl fatty acid synthase, thioesterase
C:Species: Anas platyrhynchos (mallard)
C:Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 26-May-2000
C:Accession: A00775
R:Poulsen, A.J.; Rogers, L.; Cheesbrough, T.M.; Kolattukudy, P.E.
J. Biol. Chem. 260, 15953-15958, 1985
A:Title: Cloning and sequencing of the cDNA for S-acyl fatty acid synthase thioesterase
A:Reference number: A00775; MUID:86059488; PMID:2415525
A:Accession: A00775
A:Molecule type: mRNA
A:Residues: 1-251 <POU>
A:Cross-references: GB:M12101; NID:9213098; PIDN:AAA9222.1; PID:9213099
C:Comment: This protein, isolated from the uropygial gland, catalyzes the hydrolytic re-
fatty acid synthase.
C:Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-[acyl-carrier-protein] hy-
C:Keywords: fatty acid biosynthesis; thioester hydrolase
F:17-231/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 28.0%; Score 46.5; DB 1; Length 251;
Best Local Similarity 47.8%; Pred. No. 46;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 4 SSSSFVQGFQHPDYFYMEPA 26
Db 214 SGGTSTSLPGNH---FYLMKPS 233

RESULT 43
I:50520
S:acyl fatty acid synthase thioesterase - duck
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I50520
R:Sasaki, G.C.; Cheesbrough, V.; Kolattukudy, P.E.
DNA 7, 449-457, 1988
A:Title: Nucleotide sequence of the S-acyl fatty acid synthase thioesterase gene and its
A:Reference number: I50520; MUID:89090803; PMID:2850144
A:Accession: I50520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-251 <SAS>
A:Cross-references: GB:M21635; NID:9213092; PIDN:AAA9219.1; PID:9213093
C:Genetics:
A:Insertions: 44/1; 90/2; 123/3; 180/2; 208/1
C:Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-[acyl-carrier-protein] hy-
F:17-231/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 28.0%; Score 46.5; DB 2; Length 251;
Best Local Similarity 47.8%; Pred. No. 46;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 4 SSSSFVQGFQHPDYFYMEPA 26
Db 214 SGGTSTSLPGNH---FYLMKPS 233

RESULT 44
A28550
Cyclohexanone monooxygenase (EC 1.14.13.22) - Acinetobacter sp.
C:Species: Acinetobacter sp.
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 08-Oct-1999
C:Accession: A28550; B28550
R:Chen, Y.C.J.; Peoples, O.P.; Walsh, C.T.
J. Bacteriol. 170, 781-789, 1988
A:Title: Acinetobacter cyclohexanone monooxygenase: gene cloning and sequence determinat
A:Reference number: A28550; MUID:88115180; PMID:3338974
A:Contents: NCBI 9871
A:Accession: A28550
A:Molecule type: DNA
A:Residues: 1-542 <CH1>
A:Cross-references: GB:M19029; NID:9141767; PIDN:AAA21892.1; PID:9141768
A:Accession: B28550

A:Molecule type: protein
A:Residues: 2-11 <CH2>
C:Comment: The initiator Met is not shown.
C:Keywords: oxidoreductase

Query Match 28.0%; Score 46.5; DB 2; Length 542;
Best Local Similarity 32.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

OY 2 WRSQ-SSEVQGFQHPDYFYMEP 25
Db 403 WKQPSYGVTVNNPNNFWLGP 427

RESULT 45
F82623
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82623
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82623
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-634 <SIM>
A:Cross-references: GB:AE004010; GB:AE003849; NID:99106992; PIDN:AF84709.1; GSPDB:GN
A:Experimental source: strain 945c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.
Briomes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carer
as-Neco, E.; Docena, C.; El-Dorri, H.; Facinanci, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Klatjima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, C.
Rodrigues, V.; Rosa, A.J.; de M. de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1903

Query Match 28.0%; Score 46.5; DB 2; Length 634;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 11; Conservative 6; Mismatches 8; Indels 19; Gaps 1;

OY 3 RSGSFVQGFQHPDYFYMEPA 27
Db 170 RGTERTVKTFFPITLLMFIAGVGVTYINIAQPEVLAHINPSW 213

RESULT 46
H86246
Hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H86246
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:36 ; Search time 12.4286 Seconds
(without alignments)
96.778 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78
Perfect score: 166
Sequence: 1 SWRSSGSEFVGQLFGQHPDYFLMEPAMHV 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	101	60.8	458 1	C6ST_CHICK
2	53	31.9	390 1	ARGD_METTH
3	52.5	31.6	419 1	NOE8_RHISN
4	48	28.9	240 1	MNTB_LISIN
5	48	28.9	240 1	MNTB_LISMO
6	48	28.9	338 1	GALE_MYCPN
7	48	28.9	478 1	GATA_AOUAE
8	48	28.9	902 1	GCP2_HUMAN
9	48	28.9	903 1	GCP2_MOUSE
10	48	28.9	3579 1	STAN_DROME
11	47	28.3	135 1	Y4E1_RHISN
12	47	28.3	348 1	RMLB_STRMG
13	47	28.3	465 1	CG12_CANAL
14	47	28.3	496 1	AGR2_MOUSE
15	46.5	28.0	251 1	SAST_MOUSE
16	46.5	28.0	376 1	CARD_ERWCA
17	46.5	28.0	542 1	CYMO_ACISP
18	46	27.7	234 1	ADCC_STRPN
19	46	27.7	1331 1	CYAB_LEIDO
20	45.5	27.4	423 1	AATM_CHICK
21	45.5	27.4	430 1	AATM_MOUSE
22	45.5	27.4	430 1	AATM_MOUSE
23	45.5	27.4	622 1	KDP_ECOLI
24	45.5	27.4	708 1	TRAB_HUMAN
25	45	27.1	313 1	CBR1_RHIME
26	45	27.1	313 1	CBR2_RHIME
27	45	27.1	362 1	AROC_SYNY3
28	45	27.1	381 1	OPSL_SCHGR
29	45	27.1	866 1	FIHA_HUMAN
30	45	27.1	940 1	CHRD_BRARE
31	45	27.1	1912 1	PTPD_HUMAN
32	45	27.1	2505 1	FAS_RAT
33	44.5	26.8	430 1	AATM_BOVIN
34	44.5	26.8	430 1	AATM_HUMAN
35	44.5	26.8	1396 1	VCAP_VZVD
36	44	26.5	170 1	DCP1_METMA
37	44	26.5	207 1	RCSA_SALTI
38	44	26.5	207 1	RCSA_SALTI
39	44	26.5	301 1	Y10E_ECOLI
40	44	26.5	326 1	YE09_SYNY3
41	44	26.5	351 1	ALR_DEIRA
42	44	26.5	481 1	SESL_XENLA
43	44	26.5	609 1	UVRC_HAETN
44	44	26.5	627 1	SGAB_MOUSE
45	44	26.5	627 1	SGAB_MOUSE
46	44	26.5	627 1	SGAB_RAT
47	44	26.5	632 1	SGAB_HUMAN
48	44	26.5	812 1	PEPA_HELJY
49	44	26.5	812 1	PEPA_HELJY
50	44	26.5	836 1	VGZ6_BPMD2
51	44	26.5	2105 1	POLR_ASGVP
52	43.5	26.2	2504 1	FAS_HUMAN
53	43.5	26.2	328 1	YCDU_ECOLI
54	43.5	26.2	552 1	UBIB_XYLPA
55	43	25.9	1111 1	MYSB_DICDI
56	43	25.9	153 1	GLB2_ASCSU
57	43	25.9	306 1	COSA_MYCTU
58	43	25.9	311 1	OSRP_RABIT
59	43	25.9	317 1	OK38_HUMAN
60	43	25.9	352 1	FAS_ANSAN
61	43	25.9	404 1	FTRI_YEAST
62	43	25.9	444 1	SHU5_ECOLI
63	43	25.9	445 1	YKAB_BACFI
64	43	25.9	461 1	DCOR_HUMAN
65	43	25.9	493 1	DIAG_HUMAN
66	43	25.9	511 1	MATK_HORVU
67	43	25.9	608 1	HEPAC_CAUDR
68	43	25.9	631 1	OS66_NPVLS
69	43	25.9	733 1	AGAL_PEPDE
70	43	25.9	1247 1	IRBP_HUMAN
71	43	25.9	1273 1	WEH1_YEAST
72	43	25.9	2944 1	CA17_HUMAN
73	43	25.9	3421 1	TGOU_HSVB
74	42.5	25.6	3660 1	DMD_CHICK
75	42.5	25.6	210 1	NADD_STRPY
76	42.5	25.6	338 1	CHLI_ANAPA
77	42.5	25.6	401 1	CHLI_ANAPA
78	42.5	25.6	430 1	AATM_HORSE
79	42	25.3	52 1	AATM_PIG
80	42	25.3	52 1	CRAB_COULI
81	42	25.3	174 1	CRAB_TURME
82	42	25.3	174 1	CRAB_ANAPL
83	42	25.3	225 1	CRAB_CHICK
84	42	25.3	254 1	Y339_PYRAB
85	42	25.3	254 1	MLA_VISAL
86	42	25.3	312 1	YVBG_BACSD
87	42	25.3	312 1	COSA_MYCLE
88	42	25.3	313 1	GAG_AVISN
89	42	25.3	405 1	AAT2_ARATH
90	42	25.3	405 1	AATC_DAUCA
91	42	25.3	445 1	AAT1_MEDSA
92	42	25.3	445 1	YR82_MYCLE
93	42	25.3	446 1	MGC_CHICK
94	42	25.3	449 1	AAT3_ARATH
95	42	25.3	450 1	CBPA_SYNB7
96	42	25.3	502 1	YEFA_YEAST
97	42	25.3	526 1	CAE3_DROME
98	42	25.3	718 1	CDGT_BACCI
99	42	25.3	718 1	CDGT_BACCI
100	42	25.3	902 1	FTDH_HUMAN
101	41.5	25.0	902 1	FTDH_RAT
102	41.5	25.0	189 1	HLYL_VIBPA
103	41.5	25.0	307 1	SCRK_ECOLI
104	41.5	25.0	333 1	LDHB_BRARE
105	41.5	25.0	410 1	CHLI_EUGGR
106	41.5	25.0	437 1	ATIN_VZVD
			437 1	REBH_SALTY
				P00505 homo sapien
				P09245 varicella-z
				Q4162 methanosarc
				O56083 salmonella
				O56083 salmonella
				P32667 escherichia
				P73594 synechocyst
				O97ve3 delnoccocus
				P58003 xenopus lae
				P44489 haemophilus
				P31650 mus musculu
				P31647 rattus norv
				P40066 homo sapien
				O94mv4 helicobacte
				P56070 helicobacte
				O64220 mycobacteri
				P36309 apple stem
				P49327 homo sapien
				P75910 escherichia
				O9pce8 xyfella fas
				P34092 dictyosteli
				P4672 ascaris suu
				O10807 m phosphati
				P31097 oryctolagus
				O94jx0 homo sapien
				P36189 anser anser
				P40088 saccharomyc
				P09749 escherichia
				P30268 bacillus fi
				P11926 homo sapien
				P49186 homo sapien
				P41158 hordeum vul
				Q445978 caulobacter
				P89518 leucania se
				P43467 pedicoccus
				P10745 homo sapien
				P38968 saccharomyc
				Q02388 homo sapien
				P28955 equine hepr
				P11533 gallus gall
				Q9a1f2 streptococc
				Q44498 arabidopsis
				P56571 arabidopsis
				P08907 equus caball
				P00506 sus scrofa
				O12987 columba liv
				O12995 turdus meru
				O05557 anas platyr
				O05713 gallus gall
				O9y1c7 pyrococcus
				P81446 vitiscum albu
				P37497 bacillus su
				Q9cbul m phosphati
				P03342 avian spliee
				P46645 arabidopsis
				P28734 daucus caro
				P29011 medicago sa
				O32965 mycobacteri
				P01875 gallus gall
				P46644 arabidopsis
				P39660 synechococc
				P15625 saccharomyc
				O9y192 drosophila
				P33747 bacillus ci
				O75891 homo sapien
				P28037 rattus norv
				P19249 vibrio para
				OP0v73 escherichia
				OP0v74 brachydanio
				P31205 euglena gra
				P09265 varicella-z
				P26398 salmonella

107	41.5	25.0	485	1	TRE2_SYNV3	P74130	synechocyst	180	40	24.1	454	1	AEP_YARLI	P09230	yarrowia li
108	41.5	25.0	510	1	NIFB_FRAL	P6044	frankia aln	181	40	24.1	461	1	DCOR_MOUSE	P00860	mus musculi
109	41.5	25.0	537	1	COX1_SCHPO	P07657	schizosacch	182	40	24.1	461	1	DCOR_MOUSE	P27119	mus musculi
110	41.5	25.0	622	1	COX1_BACSU	P24010	bacillus su	183	40	24.1	467	1	CIXG_LEUMC	O53080	l citrg pro
111	41.5	25.0	624	1	COX1_BACFI	P04440	bacillus fi	184	40	24.1	467	1	YAUD_RHISN	O53196	rhizobium s
112	41.5	25.0	706	1	YAEA_SCHPO	Q09880	schizosacch	185	40	24.1	468	1	CYDA_BACSU	P94364	bacillus su
113	41.5	25.0	706	1	YAEA_SCHPO	Q09880	schizosacch	185	40	24.1	468	1	CYDA_BACSU	P94364	bacillus su
114	41	24.7	145	1	CHS2_NEUCR	P30569	neurospora	186	40	24.1	473	1	REBM_SALMO	Q00473	salmonella
115	41	24.7	145	1	CHS2_NEUCR	P30569	neurospora	186	40	24.1	473	1	REBM_SALMO	Q00473	salmonella
116	41	24.7	150	1	MOAE-STACA	Q92189	staphylococ	187	40	24.1	477	1	REBM_SALMO	P24228	escherichia
117	41	24.7	150	1	MOAE-STACA	Q92189	staphylococ	187	40	24.1	477	1	REBM_SALMO	P24228	escherichia
118	41	24.7	196	1	AEQ1_AEQVI	P07164	aequorea vi	188	40	24.1	544	1	UD17_MOUSE	Q62452	mus musculi
119	41	24.7	196	1	AEQ1_AEQVI	P07164	aequorea vi	188	40	24.1	544	1	UD17_MOUSE	Q62452	mus musculi
120	41	24.7	196	1	AEQ1_AEQVI	P07164	aequorea vi	188	40	24.1	544	1	UD17_MOUSE	Q62452	mus musculi
121	41	24.7	196	1	AEQ1_AEQVI	P07164	aequorea vi	188	40	24.1	544	1	UD17_MOUSE	Q62452	mus musculi
122	41	24.7	279	1	DMSC_HAEIN	Q91117	mus musculi	189	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
123	41	24.7	279	1	DMSC_HAEIN	Q91117	mus musculi	189	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
124	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
125	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
126	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
127	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
128	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
129	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
130	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
131	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
132	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
133	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
134	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
135	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
136	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
137	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
138	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
139	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
140	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
141	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
142	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
143	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
144	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
145	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
146	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
147	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
148	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
149	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
150	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
151	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
152	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
153	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
154	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
155	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
156	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
157	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
158	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
159	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
160	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
161	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
162	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
163	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
164	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
165	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
166	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
167	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
168	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
169	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
170	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
171	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
172	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
173	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
174	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
175	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
176	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
177	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
178	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien
179	41	24.7	312	1	CDAL_YEAST	Q04204	schizosacch	190	40	24.1	663	1	RG2P_HUMAN	P54736	homo sapien

253	39	23.5	294	1	DEH1_MORSP	001398	moraxella s	326	38.5	23.2	422	1	TRE2_THEMA	094209	thermotoga
254	39	23.5	304	1	HTPX_BACDH	094966	bacillus ha	327	38.5	23.2	453	1	ARK_BRARE	042115	brachydanto
255	39	23.5	308	1	ER25_CANAL	059933	candida alb	328	38.5	23.2	472	1	C901_ARATH	042569	arabidopsis
256	39	23.5	314	1	OSTP_HUMAN	P10451	homo sapien	329	38.5	23.2	484	1	6PGD_ACTAC	P70718	actinobacil
257	39	23.5	321	1	ISP_BACCS	P29140	bacillus cl	330	38.5	23.2	489	1	6PG1_YEAST	P38720	saccharomyc
258	39	23.5	337	1	KC2A_DICD1	Q02720	dictyostell	331	38.5	23.2	497	1	YIDJ_ECOLI	P31447	escherichia
259	39	23.5	356	1	GLN1_ALINGL	004667	alnus glut	332	38.5	23.2	509	1	RUMT_DROME	P22814	dictyostell
260	39	23.5	356	1	HRPN_ERWCA	047279	erwinia car	333	38.5	23.2	510	1	GLPK_HALNI	Q9HNS	halobacteri
261	39	23.5	370	1	TPS1_HUMAN	060507	homo sapien	334	38.5	23.2	517	1	COX1_STRPU	P15544	strongyloce
262	39	23.5	370	1	TPS1_MOUSE	070281	mus musculu	335	38.5	23.2	519	1	C6G2_DROME	Q9V675	strongyloce
263	39	23.5	383	1	INSI_ECOLI	P37246	escherichia	336	38.5	23.2	566	1	GNDB_PAEIA	P23550	paenibacill
264	39	23.5	385	1	ARGD_BACSU	P36839	bacillus su	337	38.5	23.2	584	1	SACB_ACEDI	043998	acetobacter
265	39	23.5	387	1	PIGM_MOUSE	P26698	rhodococcu	338	38.5	23.2	604	1	FUCI_HAEIN	P44779	haemophilus
266	39	23.5	415	1	PAIR_MOUSE	P12388	mus musculu	339	38.5	23.2	605	1	TOPI_AERPE	Q9V601	aeropyrum p
267	39	23.5	440	1	ARCI_LYCES	P27117	bos taurus	340	38.5	23.2	673	1	MYC9_YEAST	Q9Y4X6	aeropyrum p
268	39	23.5	446	1	NRTA_SYNY3	042884	lycopersico	341	38.5	23.2	815	1	EF3_SCHPO	094489	schizosacch
269	39	23.5	450	1	DCOR_CHICK	P73452	synecocyst	342	38.5	23.2	838	1	HYA_PROMI	P16466	proteus mir
270	39	23.5	455	1	DCOR_CHICK	P27118	gallus gall	343	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
271	39	23.5	458	1	BPHA_PSEPS	P14019	cricleulius	344	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
272	39	23.5	458	1	MUC_RABIT	Q52028	pseudomonas	345	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
273	39	23.5	461	1	DCOR_BOVIN	P03988	oryctolagus	346	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
274	39	23.5	464	1	YTL3_CAEEL	P27117	bos taurus	347	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
275	39	23.5	471	1	RBL_SYNPW	010917	caenorhabdi	348	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
276	39	23.5	471	1	SYC_LISIN	P96486	synecococc	349	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
277	39	23.5	471	1	SYC_LISIN	Q92536	listeria in	350	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
278	39	23.5	473	1	YIHO_SALTY	084BD1	listeria mo	351	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
279	39	23.5	477	1	ARLY_CORGL	Q91YR4	salmonella	352	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
280	39	23.5	479	1	MUCM_RABIT	088101	corynebacte	353	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
281	39	23.5	492	1	SESI_HUMAN	P04281	oryctolagus	354	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
282	39	23.5	495	1	NIEF_KLEFX	09Y6P5	homo sapien	355	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
283	39	23.5	495	1	NIEF_KLEFX	P56567	klebsiella	356	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
284	39	23.5	507	1	YV42_CAEEL	P06772	klebsiella	357	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
285	39	23.5	511	1	YV42_CAEEL	Q18416	caenorhabdi	358	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
286	39	23.5	511	1	YV42_CAEEL	P76224	escherichia	359	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
287	39	23.5	529	1	YB53_METTH	Q27221	methanobact	360	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
288	39	23.5	531	1	FMO1_HUMAN	Q01740	homo sapien	361	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
289	39	23.5	532	1	FMO1_PIG	P16465	ratu s crofa	362	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
290	39	23.5	556	1	INNY_CAEEL	Q21123	caenorhabdi	363	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
291	39	23.5	567	1	YFZ1_SCHPO	Q9U840	schizosacch	364	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
292	39	23.5	574	1	YFZ1_SCHPO	Q9V440	dirosophila	365	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
293	39	23.5	585	1	YH70_SYNY3	Q9V440	dirosophila	366	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
294	39	23.5	623	1	EXXA_PSEAE	P73627	synecocyst	367	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
295	39	23.5	623	1	EXXA_PSEAE	Q92437	pseudomonas	368	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
296	39	23.5	647	1	YCA4_HAEIN	P44135	haemophilus	369	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
297	39	23.5	664	1	YCA4_HAEIN	P98009	acetobacter	370	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
298	39	23.5	672	1	YCA4_HAEIN	Q9Y6R2	pseudomonas	371	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
299	39	23.5	685	1	YCA4_HAEIN	P09861	dictyoglossu	372	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
300	39	23.5	698	1	YCA4_HAEIN	P09861	dictyoglossu	373	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
301	39	23.5	751	1	YCA4_HAEIN	P73627	synecocyst	374	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
302	39	23.5	751	1	YCA4_HAEIN	P73627	synecocyst	375	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
303	39	23.5	777	1	YCA4_HAEIN	P25017	agrobacteri	376	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
304	39	23.5	854	1	YCA4_HAEIN	P62178	agrobacteri	377	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
305	39	23.5	854	1	YCA4_HAEIN	O95025	homo sapien	378	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
306	39	23.5	994	1	YCA4_HAEIN	P33301	saccharomyc	379	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
307	39	23.5	1161	1	YCA4_HAEIN	Q10379	mycobacteri	380	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
308	39	23.5	1165	1	YCA4_HAEIN	P23774	simian foam	381	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
309	39	23.5	1286	1	YCA4_HAEIN	P62178	agrobacteri	382	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
310	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	383	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
311	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	384	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
312	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	385	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
313	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	386	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
314	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	387	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
315	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	388	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
316	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	389	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
317	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	390	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
318	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	391	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
319	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	392	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
320	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	393	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
321	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	394	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
322	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	395	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
323	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	396	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
324	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	397	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir
325	39	23.5	1302	1	YCA4_HAEIN	P12661	bos taurus	398	38.5	23.2	1047	1	EF3_SCHPO	P16466	proteus mir

399	38	22.9	480	1	GLGA_AGRF5	P39670	agrobacteri	472	37.5	22.6	540	1	NABR_ECOLI	P10902	escherichia
400	38	22.9	488	1	DHAL_BACST	P23239	bacillus st	473	37.5	22.6	549	1	SYBR_METTH	O26864	methanobact
401	38	22.9	492	1	SESI_MOUSE	P8006	mus musculu	474	37.5	22.6	565	1	COXI_RHOSH	P33517	rhodobacter
402	38	22.9	501	1	CP36_RABIT	P11767	oryctolagus	475	37.5	22.6	645	1	COXI_PARTE	P05489	paramecium
403	38	22.9	514	1	APPC_ECOLI	P6449	escherichia	476	37.5	22.6	722	1	VARI_HAINT	O94948	halobacteri
404	38	22.9	515	1	CD64_DROME	O9C455	m.bifunctio	477	37.5	22.6	737	1	POLG_HCVJ7	P27961	hepatitis c
405	38	22.9	527	1	PUR9_MYCLE	O92585	dictyostell	478	37.5	22.6	788	1	COXII_SULAC	O09481	sulfolobus
406	38	22.9	532	1	HEXA_DICDI	P13773	dictyostell	479	37.5	22.6	863	1	AD17_DROME	O92886	chlamydia p
407	38	22.9	534	1	FM01_RABIT	P17636	oryctolagus	480	37.5	22.6	932	1	PKNI_CHLPP	O50393	mycobacteri
408	38	22.9	545	1	ARAB_STAM	O99W76	staphylococ	481	37.5	22.6	1074	1	EMBC_MGCM	O16825	homo sapien
409	38	22.9	546	1	PUR2_AGABI	P7858	agarticus b	482	37.5	22.6	1174	1	PTNL_HUMAN	P20020	homo sapien
410	38	22.9	552	1	VNSI_EHVP2	P27585	epizootic h	483	37.5	22.6	1258	1	RPO1_ASEB7	P42486	african swi
411	38	22.9	564	1	CDCT_MOUSE	O92010	mus musculu	484	37.5	22.6	1450	1	Y645_HUMAN	O75140	homo sapien
412	38	22.9	574	1	CDCT_HUMAN	O60311	homo sapien	485	37.5	22.3	1572	1	NPRF_MOUSE	O99W48	mus musculu
413	38	22.9	584	1	V328_CHLPP	O92810	chlamydia p	486	37.5	22.3	114	1	YANA_BACLI	P30428	mus musculu
414	38	22.9	589	1	Y041_CAEEL	P34679	caenorhabdi	487	37.5	22.3	122	1	YANA_BACLI	P30428	mus musculu
415	38	22.9	635	1	TRG4_ECOLI	O00185	escherichia	488	37.5	22.3	128	1	NBSM_BOVIN	P48305	bos taurus
416	38	22.9	635	1	YDIF_BACSU	O05519	bacillus su	489	37.5	22.3	128	1	YC35_CYAPA	P48305	cyathophora
417	38	22.9	642	1	DCHS_MOUSE	O05519	bacillus su	490	37.5	22.3	146	1	BAHG_VITST	P04452	vitreoscilli
418	38	22.9	662	1	COL4_BRAJA	P23738	mus musculu	491	37.5	22.3	147	1	GLB1_ARTSX	P19363	artemia sp.
419	38	22.9	666	1	LTP_STARP	P88057	bradyrhizob	492	37.5	22.3	178	1	YMO6_PARTE	P15607	paramecium
420	38	22.9	694	1	MALQ_ECOLI	O02510	staphylococ	493	37.5	22.3	194	1	R1SE_PYRAB	O9V0D2	pyrococcus
421	38	22.9	716	1	ENPL_RABIT	P15977	escherichia	494	37.5	22.3	199	1	YPRN_BACSU	P50741	bacillus su
422	38	22.9	750	1	PSAA_MARPO	O18750	oryctolagus	495	37.5	22.3	203	1	HPR_BACSU	P11065	bacillus su
423	38	22.9	750	1	PSAA_PSINU	P64406	marichantia	496	37.5	22.3	208	1	YCAC_ECOLI	P1267	escherichia
424	38	22.9	751	1	PSAA_EUGGR	O60432	psilotum nu	497	37.5	22.3	209	1	GTP_CRILLO	P26424	cricetulus
425	38	22.9	752	1	METE_ECOLI	P25665	escherichia	498	37.5	22.3	209	1	GTP_HUMAN	P09211	homo sapien
426	38	22.9	753	1	METE_SALT	O91661	salmonella	499	37.5	22.3	209	1	GTP_MACMU	O28514	mascaca mula
427	38	22.9	795	1	ENPL_CHICK	P08110	gallus gall	500	37.5	22.3	213	1	Y014_METUA	P66887	mycobacteri
428	38	22.9	802	1	ENPL_MOUSE	P08113	mus musculu	501	37.5	22.3	222	1	Y014_METUA	P66887	mycobacteri
429	38	22.9	803	1	ENPL_HUMAN	P14623	homo sapien	502	37.5	22.3	227	1	DRNG_AERHY	O31222	aeromonas h
430	38	22.9	804	1	ENPL_CANFA	P41148	canis famli	503	37.5	22.3	268	1	LGT_HABIN	P44630	haemophilus
431	38	22.9	804	1	YRR2_CAEEL	O09392	sus scrofa	504	37.5	22.3	273	1	AROE_BUCAI	O44607	buchnera ap
432	38	22.9	823	1	POLG_HRVIA	P26446	gallus gall	505	37.5	22.3	277	1	SP0A_CLOIN	P52933	clostridium
433	38	22.9	832	1	POLG_HRVIA	O15909	typanosoma	506	37.5	22.3	280	1	LPXA_CHLTR	O84536	chlamydia t
434	38	22.9	838	1	RTR1_TRYBB	O92398	chlamydia p	507	37.5	22.3	281	1	LGT_BUCAI	P57514	buchnera ap
435	38	22.9	928	1	PMP9_CHLPP	P26446	gallus gall	508	37.5	22.3	286	1	YNB3_VIBCH	O9K14	vibrio chol
436	38	22.9	950	1	DC12_DROME	P26446	gallus gall	509	37.5	22.3	291	1	KRPP_ALCEU	P193924	alcaligenes
437	38	22.9	1011	1	POLG_CHICK	P26446	gallus gall	510	37.5	22.3	303	1	PANE_SAVTI	O82853	salmonella
438	38	22.9	1037	1	CAR6_HUMAN	O9K69	homo sapien	511	37.5	22.3	310	1	LACC_LACVI	P23391	lactococcus
439	38	22.9	1067	1	CARB_DROME	O8K843	clostridium	512	37.5	22.3	311	1	PYDB_LACIA	O9C748	lactococcus
440	38	22.9	1123	1	DC11_DROME	P18169	drosophila	513	37.5	22.3	317	1	GPDA_SYN3	P73033	synchocyst
441	38	22.9	1202	1	UNG2_RAT	P03179	rattus norv	514	37.5	22.3	319	1	YB98_MYCPN	O50290	mycoplasma
442	38	22.9	1318	1	VP14_EBV	P03179	epstein-bar	515	37.5	22.3	321	1	BLAC_BACVU	P30899	bacteroides
443	38	22.9	1377	1	V198_HUMAN	O9UBW7	homo sapien	516	37.5	22.3	322	1	PPI6_ARATH	P48486	arabidopsis
444	38	22.9	1380	1	CYAA_LEIDO	O27675	leishmania	517	37.5	22.3	331	1	COL6_ARATH	O91427	arabidopsis
445	38	22.9	1465	1	MYM2_HUMAN	P54296	homo sapien	518	37.5	22.3	335	1	MAER_CLOBU	O05103	clostridium
446	38	22.9	1589	1	DC13_DROME	P13368	drosophila	519	37.5	22.3	340	1	OMP2_CHLTR	O9P113	chlamydia m
447	38	22.9	1681	1	CTH_CAEEL	P34574	caenorhabdi	520	37.5	22.3	340	1	OMP2_CHLTR	P38006	chlamydia t
448	38	22.9	1695	1	KR1A_MOUSE	P33173	mus musculu	521	37.5	22.3	342	1	GCP_HAEIN	P43764	haemophilus
449	38	22.9	2157	1	POLG_HRVIA	P13916	human rhino	522	37.5	22.3	343	1	REP7_ECOLI	O52221	escherichia
450	38	22.9	2157	1	SPCB_DROME	O00963	drosophila	523	37.5	22.3	343	1	REP7_ECOLI	P18023	escherichia
451	38	22.9	2291	1	POLN_ONNNG	P13886	c.nyong-nyo	524	37.5	22.3	344	1	OMP2_CHLPP	O92752	chlamydia p
452	38	22.9	2514	1	7LES_DROME	P13368	drosophila	525	37.5	22.3	345	1	HK16_LYCES	O22299	lycopersico
453	38	22.6	151	1	EPRI_THEAC	O9LKM3	thermoplas	526	37.5	22.3	356	1	P2C2_CAEEL	P49596	caenorhabdi
454	37.5	22.6	187	1	EPF_MYCTU	P95019	mycobacteri	527	37.5	22.3	357	1	YQO2_CAEEL	O09305	caenorhabdi
455	37.5	22.6	226	1	NARV_ECOLI	P19316	escherichia	528	37.5	22.3	362	1	CARA_BACHD	O9K9V8	chlamydia ha
456	37.5	22.6	241	1	AGL8_SINAL	O41274	sinapis alb	529	37.5	22.3	368	1	RTL_CHIRE	P11660	chlamydomon
457	37.5	22.6	241	1	VMOM_BPVU	P06018	bacterioph	530	37.5	22.3	376	1	OPSD_SRRSP	P35362	sphodromant
458	37.5	22.6	275	1	HXDC_HUMAN	P35452	homo sapien	531	37.5	22.3	377	1	VEZ_HPV13	O02263	human papil
459	37.5	22.6	278	1	RHAS_SALT	P27029	salmonella	532	37.5	22.3	385	1	ARGD_THEMA	O9X245	thermopoga
460	37.5	22.6	279	1	HXDC_MOUSE	P23812	mus musculu	533	37.5	22.3	388	1	ARGD_MENJA	O58131	methanococ
461	37.5	22.6	286	1	AROE_LACIA	O9CE57	lactococcus	534	37.5	22.3	400	1	RUN5_BACAG	O85465	bacillus ag
462	37.5	22.6	313	1	LDH_LISIN	O9CF65	listeria in	535	37.5	22.3	400	1	KAP2_MOUSE	P12367	mus musculu
463	37.5	22.6	313	1	LDH_LISIN	P33380	listeria in	536	37.5	22.3	401	1	ALKA_PSPOL	P12367	mus musculu
464	37.5	22.6	329	1	YX99_BACHD	P53380	listeria ha	537	37.5	22.3	402	1	KROS_AYASU	P00529	avian sarco
465	37.5	22.6	345	1	DMFG_PSESP	P51016	pseudomonas	538	37.5	22.3	403	1	PRSE_DICDI	P34123	dictyostell
466	37.5	22.6	435	1	DMG3_RHIME	O87393	rhizobium m	539	37.5	22.3	407	1	HEM1_RHOSH	O04512	rhodobacter
467	37.5	22.6	458	1	PSHC_PROMO	P51753	prochloroth	540	37.5	22.3	409	1	GUN2_BACSA	P06565	bacillus sp
468	37.5	22.6	475	1	IFZG_DROME	O24208	drosophila	541	37.5	22.3	410	1	ILBC_FELCA	O9M266	felis silve
469	37.5	22.6	516	1	GIC5_SOYBN	P04347	glycine max	542	37.5	22.3	413	1	PRSE_SOLTU	P54778	solanum tub
470	37.5	22.6	529	1	TYRO_CHICK	P33024	gallus gall	543	37.5	22.3	413	1	YBSA_YEAST	P38244	saccharomyc
471	37.5	22.6	537	1	YK00_MYCTU	O10857	mycobacteri	544	37.5	22.3	414	1	PRSE_CAEEL	P46502	caenorhabdi

545	37	22.3	415	1	PRSM6_MANE	P45507	manduca sex	618	37	22.3	1980	1	MY9B_RAT	O63358	rattus norv
546	37	22.3	417	1	DSRA_CHRV1	O33998	chromatium	619	37	22.3	2096	1	BP28_DROME	O63525	drosophila
547	37	22.3	418	1	PSR6_HUMAN	P43686	homo sapien	620	37	22.3	2109	1	RRLP_VSVet	O35723	vesicular s
548	37	22.3	418	1	PRSM_MOUSE	P54775	mus musculu	621	37	22.3	2114	1	MY9B_MOUSE	O63906	mus musculu
549	37	22.3	418	1	PRSM_RAT	O63570	rattus norv	622	37	22.3	2224	1	FAS_HUMAN	P12258	homo sapien
550	37	22.3	426	1	P967_CONGL	P33865	cornebacte	623	37	22.3	2388	1	SPCF_RAT	P12258	homo sapien
551	37	22.3	430	1	SUN_CONHL	P45677	coxiella bu	624	37	22.3	2390	1	SPCF_HUMAN	O15020	homo sapien
552	37	22.3	432	1	VG3L_YEAST	P48236	saccharomyc	625	37	22.3	4349	1	DYHC_FUSO	O34748	schistosom
553	37	22.3	438	1	Y82_MTCU	O33324	mycobacteri	626	37	22.0	89	1	ISP2_BRAVA	P18716	fusarium so
554	37	22.3	439	1	ARP3_NEOCR	P178712	neutrospora	627	37	22.0	190	1	GLPG_HAETN	P30961	bradyrhizob
555	37	22.3	440	1	YA33_SCHPO	O09710	schizosacch	628	37	22.0	192	1	GLPG_HAETN	P30961	bradyrhizob
556	37	22.3	453	1	SLY4_YEAST	P22215	saccharomyc	629	37	22.0	207	1	NADD_HAETN	P44783	haemophilus
557	37	22.3	455	1	MOC_MOUSE	P11872	mus musculu	630	37	22.0	254	1	NADD_BACHD	O4783	haemophilus
558	37	22.3	457	1	ODR7_CAEBL	P14933	caenorhabdi	631	37	22.0	258	1	PDHR_ECOLI	O60957	escherichia
559	37	22.3	467	1	VP25_CAEBL	O09436	caenorhabdi	632	37	22.0	283	1	CAFM_YERE	P28928	yersinia pe
560	37	22.3	470	1	VP25_CAEBL	O09436	caenorhabdi	633	37	22.0	286	1	HTPX_MTCU	P70266	mus musculu
561	37	22.3	474	1	RBL_PSEH	O11179	caenorhabdi	634	37	22.0	328	1	HOLB_PSEAE	O06429	mycobacteri
562	37	22.3	476	1	MOC_MOUSE	P01873	mus musculu	635	37	22.0	334	1	IDHB_FUNNE	P20371	fundulus be
563	37	22.3	482	1	CATA_PSEAE	O52762	pseudomonas	636	37	22.0	351	1	HM14_CAEBL	O20277	fundulus be
564	37	22.3	483	1	KPKY_METEX	O05118	methylobact	637	37	22.0	371	1	OPSL_CALVT	P22269	calliphora
565	37	22.3	488	1	GUN1_BACS4	P06556	bacillus sp	638	37	22.0	373	1	MATC_NEOCR	O10115	neutrospora
566	37	22.3	488	1	SSDH_RAT	P11650	rattus norv	639	37	22.0	376	1	TPS2_MOUSE	O68856	mus musculu
567	37	22.3	491	1	HMCS_YEAST	P34839	saccharomyc	640	37	22.0	377	1	TPS2_HUMAN	O67070	homo sapien
568	37	22.3	507	1	KROS_CHICK	P08941	gallus galli	641	37	22.0	419	1	FRAA_HUMAN	P16930	homo sapien
569	37	22.3	508	1	C6D5_DROME	O9yifp1	drosophila	642	37	22.0	421	1	NADR_HAETN	P44308	haemophilus
570	37	22.3	528	1	COX1_TIRIU	O01555	trichophyto	643	37	22.0	432	1	SYN_LACBE	P54262	lactobacilli
571	37	22.3	534	1	COX1_MLTYA	P20386	kluyeromyc	644	37	22.0	446	1	ATPD_MTCLE	P53006	mycobacteri
572	37	22.3	541	1	YB36_METUA	O68536	methanococc	645	37	22.0	473	1	PSBC_SECE	P10804	secale cere
573	37	22.3	542	1	MATK_ORISA	P12175	oryza sativ	646	37	22.0	475	1	EX1_ECOLI	O4995	escherichia
574	37	22.3	542	1	MATK_ORISA	P12175	oryza sativ	647							

691	36	21.7	165	1	YRH1_RHOSH	053229 rhodobacter	764	36	21.7	446	1	Y612_METUA	Q58029 methanococ
692	36	21.7	186	1	KGUA_YEAST	P15454 saccharomyc	765	36	21.7	450	1	CSK_HUMAN	P41840 homo sapien
693	36	21.7	192	1	BM3R_BACME	P33506 bacillus me	766	36	21.7	453	1	NRAM_IAMIL	P03470 influenza a
694	36	21.7	194	1	RISE_PIRHO	Q87706 pyrococcus	767	36	21.7	465	1	EXSH_RHIME	P03680 rhizobium m
695	36	21.7	195	1	PAPH_ECOLI	P07111 escherichia	768	36	21.7	469	1	NRAM_TIPAR	P06619 influenza a
696	36	21.7	195	1	PSRH_ECOLI	P2185 escherichia	769	36	21.7	472	1	SYC_CHLMO	P23621 neisseria m
697	36	21.7	199	1	KGUA_ANASPI	Q82017 anabaena sp	770	36	21.7	475	1	SYC_CHLMO	Q84787 chlamydia m
698	36	21.7	208	1	ACPD_STRAM	Q94911 streptomyces	771	36	21.7	477	1	YF45_AOUAE	Q67495 aquifex aeo
699	36	21.7	216	1	PAAD_STRGO	Q92617 chlamydia p	772	36	21.7	479	1	REFM_SALTY	P26404 salmonella
700	36	21.7	219	1	BROD_CHLUP	Q92617 chlamydia p	773	36	21.7	481	1	Y772_MYCTU	O10875 mycobacteri
701	36	21.7	220	1	KCY_LACIA	Q92617 lactococcus	774	36	21.7	484	1	CATP_HUMAN	Q94911 homo sapien
702	36	21.7	221	1	KC2B_SPOFR	P26483 spodioplera	775	36	21.7	485	1	VAIB_TRYCO	Q26376 trypanosoma
703	36	21.7	234	1	KC2B_CAEEL	P28548 caenorhabdi	776	36	21.7	491	1	CPB3_RAT	P13107 rattus norv
704	36	21.7	246	1	DCMH_BPT2	P18028 bacterioph	777	36	21.7	491	1	CPB3_RAT	P20813 homo sapien
705	36	21.7	248	1	EX15_ARATH	Q80622 arabidopsis	778	36	21.7	493	1	CPB6_HUMAN	P46371 rhodococcus
706	36	21.7	249	1	NODH_RHITH	P22994 rhizobium t	779	36	21.7	493	1	YTR2_RHOER	Q35074 mus musculu
707	36	21.7	255	1	RS2_LACIA	Q92617 lactococcus	780	36	21.7	501	1	PTGI_MOUSE	P51338 rattus norv
708	36	21.7	255	1	EXPE_ARATH	Q98865 arabidopsis	781	36	21.7	503	1	CP33_RAT	Q59921 emeritella
709	36	21.7	263	1	YCS6_PORPU	P21208 porphyra pu	782	36	21.7	504	1	TPS1_EMDNI	P51338 rattus norv
710	36	21.7	270	1	KITM_MOUSE	Q91088 mus musculu	783	36	21.7	505	1	MAUD_STNY3	P72785 synechocyst
711	36	21.7	277	1	DGX_MOUSE	Q94911 s peptid m	784	36	21.7	512	1	Y4DC_RHISN	Q53197 rhizobium s
712	36	21.7	290	1	LEP_HELPU	P35593 s peptid m	785	36	21.7	515	1	CPX1_ONCMY	P78968 schizosacch
713	36	21.7	291	1	KPRR_ALCEU	P19923 alcaligenes	786	36	21.7	516	1	TPSA_ASPNG	P48170 oncorhynch
714	36	21.7	293	1	CC28_YEAST	P73548 s phosphati	787	36	21.7	517	1	YFSA_AZQVI	Q00075 aspergillu
715	36	21.7	298	1	SB81_HUMAN	P00546 saccharomyc	788	36	21.7	521	1	CYOA_HAEIN	P45021 haemophilu
716	36	21.7	298	1	YIGM_ECOLI	O00566 homo sapien	789	36	21.7	522	1	YTOE_YEAST	P11026 escherichia
717	36	21.7	299	1	ER25_YEAST	P27849 escherichia	790	36	21.7	529	1	YTOE_YEAST	P40458 saccharomyc
718	36	21.7	309	1	MSAB_STREY	P31045 saccharomyc	791	36	21.7	531	1	YAGH_ECOLI	Q94911 drosophila
719	36	21.7	309	1	MSAB_STREY	Q95911 s peptid m	792	36	21.7	536	1	YAGH_ECOLI	P77713 escherichia
720	36	21.7	312	1	MAB1_STRPN	P35593 s peptid m	793	36	21.7	537	1	CYDA_AZQVI	Q09049 azotobacter
721	36	21.7	313	1	GSBH_RHITO	Q98068 rhizobium l	794	36	21.7	587	1	GOX_PENAG	P81156 penicillium
722	36	21.7	315	1	GSBH_RHIME	Q92617 rhizobium m	795	36	21.7	588	1	CYDD_ECOLI	P29018 escherichia
723	36	21.7	319	1	GP31_HUMAN	Q00270 homo sapien	796	36	21.7	589	1	SYRF_MOUSE	Q94911 mus musculu
724	36	21.7	324	1	YPDA_BACSU	P50736 bacillus su	797	36	21.7	605	1	GOX_TALEFL	Q92452 talaromyces
725	36	21.7	330	1	ARGC_PYRAB	Q94911 pyrococcus	798	36	21.7	610	1	FIMB_DICDI	P54580 dictyostell
726	36	21.7	330	1	ARGC_PYRAB	O53397 pyrococcus	799	36	21.7	622	1	DCTB_RHILE	P10066 xentopus lae
727	36	21.7	331	1	LDHA_TRASC	Q92617 tracheomys s	800	36	21.7	629	1	K6AB_XENLA	P10666 xentopus lae
728	36	21.7	333	1	Y642_AMEBE	O58059 methanococ	801	36	21.7	632	1	VEI_HPV31	P17382 human papil
729	36	21.7	333	1	KC1B_BOVIN	P33507 bos taurus	802	36	21.7	632	1	Y219_HUMAN	Q94911 mus musculu
730	36	21.7	336	1	LG6C_HUMAN	Q94911 mus musculu	803	36	21.7	633	1	Y219_HUMAN	P18654 mus musculu
731	36	21.7	336	1	NOSO_BACSU	O34453 bacillus su	804	36	21.7	642	1	FIMB_YEAST	P32599 saccharomyc
732	36	21.7	341	1	Y6XE_BACSU	P31964 bacillus su	805	36	21.7	642	1	SYRM_YEAST	P38714 saccharomyc
733	36	21.7	343	1	SP2D_BACSU	P07372 bacillus su	806	36	21.7	643	1	PLBI_TORDE	O11121 torulaspora
734	36	21.7	343	1	UPR2_MANES	Q40285 manihot esc	807	36	21.7	649	1	KR74_HSV11	Q00095 ictaluriid h
735	36	21.7	345	1	GLN1_SOYBN	P24089 glycine max	808	36	21.7	673	1	NCPR_PHAU1	P37116 phaseolus a
736	36	21.7	355	1	GOP2_CAEEL	P46577 caenorhabdi	809	36	21.7	689	1	PSA1_PROHO	Q94911 mus musculu
737	36	21.7	358	1	TRMU_THEMA	Q94911 thermotoga	810	36	21.7	716	1	HS83_DROAV	O02192 drosophila
738	36	21.7	361	1	COL9_ARATH	O22800 arabidopsis	811	36	21.7	717	1	HS83_DROME	P02628 drosophila
739	36	21.7	364	1	CARA_BACCL	P53482 pseudomonas	812	36	21.7	719	1	AGAL_PEDPE	P43469 mus musculu
740	36	21.7	368	1	PPBL_PSEAE	O60564 mesocricetu	813	36	21.7	724	1	K6A1_MOUSE	P18653 mus musculu
741	36	21.7	369	1	LKXB_MESAU	P54111 kluyveromyc	814	36	21.7	725	1	CTPC_MYCLE	Q96474 brachydanio
742	36	21.7	370	1	GAA2_KLUTIA	O88609 mus musculu	815	36	21.7	726	1	HS9A_BRARE	O02942 blaberus di
743	36	21.7	372	1	LKXB_MOUSE	Q92617 bos taurus	816	36	21.7	726	1	TRF_BIADI	P20491 pseudomonas
744	36	21.7	373	1	CS14_BOVIN	P06002 drosophila	817	36	21.7	731	1	K6A2_HUMAN	O15349 homo sapien
745	36	21.7	373	1	OPSI_DROME	O91498 pseudomonas	818	36	21.7	733	1	K6A2_MOUSE	P10665 mus musculu
746	36	21.7	378	1	LPXB_PSEAE	O60663 homo sapien	819	36	21.7	733	1	K6A2_MOUSE	O15418 homo sapien
747	36	21.7	379	1	TRMU_XYLFA	Q70081 caenorhabdi	820	36	21.7	735	1	K6A1_HUMAN	O63531 rattus norv
748	36	21.7	380	1	TEPSA_CAEEL	O94911 mus musculu	821	36	21.7	740	1	K6A3_HUMAN	O15111 h inhibitor
749	36	21.7	382	1	FEIB_HUMAN	Q94911 mus musculu	822	36	21.7	745	1	IKKA_HUMAN	O60680 mus musculu
750	36	21.7	382	1	FEIB_HUMAN	Q94911 mus musculu	823	36	21.7	745	1	IKKA_HUMAN	O60680 mus musculu
751	36	21.7	387	1	MER2_CAEEL	P48290 leptospira	824	36	21.7	745	1	IKKA_HUMAN	O60680 mus musculu
752	36	21.7	387	1	MER2_CAEEL	P48290 leptospira	825	36	21.7	745	1	IKKA_HUMAN	O60680 mus musculu
753	36	21.7	388	1	UDG_SALTY	O04873 mycobacteri	826	36	21.7	752	1	K6A6_HUMAN	O94911 mus musculu
754	36	21.7	395	1	DNAT_MYCTU	O07703 bordetella	827	36	21.7	752	1	K6A6_HUMAN	O94911 mus musculu
755	36	21.7	395	1	DNAT_MYCTU	O07703 bordetella	828	36	21.7	753	1	PSAA_PINTH	P41639 pius thump
756	36	21.7	397	1	YAK7_SCHPO	Q92617 schizosacch	829	36	21.7	753	1	FTSH_MYCTU	P96942 mycobacteri
757	36	21.7	397	1	YAK7_SCHPO	Q92617 schizosacch	830	36	21.7	753	1	FTSH_MYCTU	P96942 mycobacteri
758	36	21.7	399	1	PKNA_STRPN	Q92617 streptococ	831	36	21.7	759	1	RS64_MOUSE	P75382 mycoplasma
759	36	21.7	403	1	YDHQ_ECOLI	P37597 escherichia	832	36	21.7	799	1	RS64_MOUSE	Q92268 mus musculu
760	36	21.7	423	1	PRSG_ASPNG	P78578 aspergillus	833	36	21.7	804	1	YB8P_ECOLI	E77504 escherichia
761	36	21.7	428	1	HGD_DROME	Q94911 drosophila	834	36	21.7	812	1	RIR1_DROME	P48991 drosophila
762	36	21.7	434	1	KES1_YEAST	P35844 saccharomyc	835	36	21.7	847	1	MDOH_ECOLI	P33337 escherichia
763	36	21.7	445	1	ESAL_YEAST	Q08649 saccharomyc	836	36	21.7	867	1	PMPI_CHLNU	Q94911 chlamydia m

837	36	21.7	874	1	CLPA_BRANA	P46523	brasica na	910	35.5	21.4	409	1	GP19_MOUSE	Q61121	mus musculus
838	36	21.7	877	1	INCE_SACKI	P53552	gallus gall	911	35.5	21.4	413	1	OR7A_DROME	Q9w315	drosophila
839	36	21.7	880	1	BG12-CHICK	P22507	saccharomyc	912	35.5	21.4	424	1	OR7A_DROME	P29620	oryza sativa
840	36	21.7	892	1	HUL4_YEAST	P40985	saccharomyc	913	35.5	21.4	426	1	FIBG_ADE04	P36844	human adeno
841	36	21.7	898	1	ACOC_CUCMA	P49608	cucurbita m	914	35.5	21.4	432	1	FIBG_PETMA	P04115	petromyzon
842	36	21.7	910	1	PMW2_CAEEL	P91441	caenorhabdi	915	35.5	21.4	435	1	GLN3_RHILP	P31592	rhizobium
843	36	21.7	911	1	NIAT_BRANA	P39867	brassica na	916	35.5	21.4	440	1	YD4C_SCHPO	Q10305	schizosacch
844	36	21.7	922	1	CIRA_PPA	P35000	pisum sativ	917	35.5	21.4	445	1	FIBG_RAT	P02680	ratius norv
845	36	21.7	923	1	CLAB_LYCES	P31542	lycopersico	918	35.5	21.4	449	1	TRB2_AERPE	Q9y9h2	aeropyrum p
846	36	21.7	926	1	CLAA_LYCES	P31541	lycopersico	919	35.5	21.4	469	1	APTE_DROME	P29673	drosophila
847	36	21.7	974	1	SIN4_YEAST	P32259	saccharomyc	920	35.5	21.4	471	1	SYC_SUITO	Q969C3	sulfolobus
848	36	21.7	977	1	YD68_SCHPO	P10495	schizosacch	921	35.5	21.4	495	1	ION2_THENO	P56275	thermoplas
849	36	21.7	1002	1	EPB5_CHICK	Q07497	gallus gall	922	35.5	21.4	507	1	EPOR_MOUSE	P14753	mus musculu
850	36	21.7	1005	1	MANA_DICDI	P34098	dactyostell	923	35.5	21.4	508	1	YBCK_ECOLI	P77698	escherichia
851	36	21.7	1008	1	VGIM_HUK	P09613	unikonteml v	924	35.5	21.4	512	1	COX1_LOCM1	Q36421	locusta mig
852	36	21.7	1009	1	M2B2_HUMAN	Q9y4e5	homo sapien	925	35.5	21.4	513	1	COX1_YEAST	P00401	saccharomyc
853	36	21.7	1018	1	M2B2_MOUSE	O54782	mus musculu	926	35.5	21.4	515	1	COX1_ORNAN	Q36452	ornithorhyn
854	36	21.7	1022	1	DPO5_YEAST	P39985	saccharomyc	927	35.5	21.4	516	1	COX1_PELST	P00397	mus musculu
855	36	21.7	1026	1	VG37_BEPT4	P03744	bacteriophag	928	35.5	21.4	514	1	COX1_MOUSE	O78672	mus musculu
856	36	21.7	1032	1	YA60_SYNY3	P72637	synecocyst	929	35.5	21.4	514	1	COX1_PROM1	Q05143	pelotomeda
857	36	21.7	1039	1	ITAB_HUMAN	P08514	homo sapien	930	35.5	21.4	514	1	COX1_RAT	Q05103	ratius norv
858	36	21.7	1087	1	PER_DROVI	P12249	drosophila	931	35.5	21.4	516	1	COX1_BRARE	Q9m1y8	brachydanio
859	36	21.7	1097	1	DPO1_MOMVS	P27172	murine cyto	932	35.5	21.4	516	1	COX1_CARAU	Q78681	carassius a
860	36	21.7	1129	1	RPA2_DROME	P20028	drosophila	933	35.5	21.4	516	1	COX1_CROLA	P34188	crossostoma
861	36	21.7	1135	1	PHYC_SORBI	P23528	sorghum bic	934	35.5	21.4	516	1	COX1_CYPCA	P24395	cyprinus ca
862	36	21.7	1143	1	DP2L_ARCFU	O28552	archaeoglob	935	35.5	21.4	517	1	COX1_STRCA	Q21399	struthio ca
863	36	21.7	1157	1	POL_SFV3L	P27401	simian foam	936	35.5	21.4	517	1	COX1_MYGXL	Q21079	myxine glut
864	36	21.7	1331	1	MANB_GALSA	P22533	caldocellum	937	35.5	21.4	517	1	COX1_PETMA	P24794	beyza vulgar
865	36	21.7	1333	1	VGRI_MOUSE	P35669	mus musculu	938	35.5	21.4	522	1	COX1_MARPO	P26856	marchantia
866	36	21.7	1336	1	VGRI_RAT	P53767	ratius norv	939	35.5	21.4	524	1	COX1_BETVU	P14578	oryza sativ
867	36	21.7	1338	1	VGRI_HUMAN	P17948	h vascular	940	35.5	21.4	524	1	COX1_ORYSA	P14578	oryza sativ
868	36	21.7	1350	1	NOS_DROME	Q27571	drosophila	941	35.5	21.4	527	1	COX1_MHEAT	P068741	triticum ae
869	36	21.7	1364	1	PUR2_DROPS	P16340	d trifuncti	942	35.5	21.4	527	1	COX1_ARATH	Q07063	arabidopsis
870	36	21.7	1376	1	VCAP_HSEVB	P27190	equine herp	943	35.5	21.4	528	1	COX1_MATZE	P05742	zea mays (m
871	36	21.7	1448	1	TRC5_ECOLI	P42695	homo sapien	944	35.5	21.4	529	1	TYRO_HUMAN	P14679	homo sapien
872	36	21.7	1507	1	Y056_HUMAN	P24985	homo sapien	945	35.5	21.4	530	1	COX1_SORBI	P05502	sorghum bic
873	36	21.7	1509	1	PDRI_YEAST	Q04182	saccharomyc	946	35.5	21.4	533	1	COX1_DINSE	O75548	dinodone sem
874	36	21.7	1529	1	BIME_YEAST	P24686	emeritella	947	35.5	21.4	534	1	TYRO_MOUSE	P11344	mus musculu
875	36	21.7	2033	1	SPCB_MOUSE	P15508	mus musculu	948	35.5	21.4	534	1	COX1_SACDO	P98001	saccharomyc
876	36	21.7	2137	1	SPCB_MOUSE	P11377	homo sapien	949	35.5	21.4	538	1	COX1_PSEAE	O51363	pseudomonas
877	36	21.7	2175	1	POLG_BOVEY	P128107	bos taurus	950	35.5	21.4	554	1	CXIA_PARDE	P03305	paracoccus
878	36	21.7	2211	1	FAS_BOVIN	O28107	bos taurus	951	35.5	21.4	557	1	COX1_NEUCR	P03943	neurospora
879	36	21.7	2258	1	FAS_PIG	O991P1	sus scrofa	952	35.5	21.4	567	1	COX1_EMENT	P80C_HUMAN	homo sapien
880	36	21.7	2278	1	FAPL_YEAST	P34756	saccharomyc	953	35.5	21.4	576	1	AYMW_BACLI	Q04977	bacillus li
881	36	21.7	2875	1	RRL1_TSW1	P28976	tomato spot	954	35.5	21.4	578	1	LMRA_LACIA	C49A_DROME	lacococcus
882	35.5	21.4	46	1	LYTA_CHLTR	Q94603	chlamydia t	955	35.5	21.4	584	1	COXN_BRATA	P98000	bradyrhizob
883	35.5	21.4	150	1	YV35_PSEAE	Q9hyh1	pseudomonas	956	35.5	21.4	591	1	COXN_PHYPO	Q07434	physarum po
884	35.5	21.4	172	1	COX1_EPHEO	O33439	ephedra equ	957	35.5	21.4	594	1	WD66_PHYPO	P90587	physarum po
885	35.5	21.4	187	1	EPF_MYCLE	O9csc0	mycobacteri	958	35.5	21.4	601	1	WDRI_HUMAN	O75083	homo sapien
886	35.5	21.4	189	1	HIV2_VIBPA	P19250	vibrio para	959	35.5	21.4	606	1	PEPA_RICCO	O92327	gallus gall
887	35.5	21.4	209	1	LIPB_RICCN	O921X0	rickettsia	960	35.5	21.4	609	1	COX1_CHICK	Q41140	ricinus com
888	35.5	21.4	254	1	PDHR_SALTY	O911h9	salmomella	961	35.5	21.4	617	1	PEPA_RICCO	Q41140	ricinus com
889	35.5	21.4	254	1	TRYP_SARBU	P51588	sarcophaga	962	35.5	21.4	617	1	PEPA_RICCO	Q41140	ricinus com
890	35.5	21.4	274	1	RT02_PROM1	P46741	protoltheca	963	35.5	21.4	625	1	YBAG_YEAST	P38750	saccharomyc
891	35.5	21.4	282	1	ICE3_XENLA	P55666	xenopus lae	964	35.5	21.4	627	1	YBAG_YEAST	P38750	saccharomyc
892	35.5	21.4	297	1	ATP6_MYCGA	P33351	mycoplasma	965	35.5	21.4	633	1	BZ2L_YEAST	P38822	saccharomyc
893	35.5	21.4	302	1	NAMC_PSEAE	O57145	pseudomonas	966	35.5	21.4	636	1	PAH1_HUMAN	P11940	homo sapien
894	35.5	21.4	303	1	DEOC_CAEEL	Q19264	caenorhabdi	967	35.5	21.4	636	1	CTO1_SOUAC	O73925	squalus aca
895	35.5	21.4	305	1	FISX_NITGO	P91567	neisseria g	968	35.5	21.4	660	1	YBAG_YEAST	P18773	schizosacch
896	35.5	21.4	321	1	YECF_HAEIN	P44167	haemophilus	969	35.5	21.4	662	1	YBAG_YEAST	P18773	schizosacch
897	35.5	21.4	323	1	IDH_LACHE	O32765	lactobacill	970	35.5	21.4	664	1	YBAG_YEAST	P18773	schizosacch
898	35.5	21.4	324	1	ARG1_AGRFS	P14112	agrobacteri	971	35.5	21.4	688	1	YBAG_YEAST	P18773	schizosacch
899	35.5	21.4	327	1	XVNA_ASPAC	O59859	aspergillus	972	35.5	21.4	825	1	VGRI_HAEIN	P18773	schizosacch
900	35.5	21.4	330	1	GPTR_PYRPO	O05991	pyrus commu	973	35.5	21.4	841	1	VGRI_VZVD	P18773	schizosacch
901	35.5	21.4	337	1	COX1_NORPE	O03359	linum maj	974	35.5	21.4	890	1	STVL_YEAST	P37296	saccharomyc
902	35.5	21.4	337	1	COX1_TINMA	O03554	ratius norv	975	35.5	21.4	904	1	YBAG_YEAST	P37296	saccharomyc
903	35.5	21.4	352	1	GP19_RAT	P70585	synecocyst	976	35.5	21.4	928	1	CHS2_EXODE	P30601	exophiala d
904	35.5	21.4	357	1	CHLI_SYNY3	P51634	synecocyst	977	35.5	21.4	983	1	ENV_VILIV	P30379	visna lent
905	35.5	21.4	361	1	DP43_RHIO	O984W7	rhizobium l	978	35.5	21.4	993	1	SMB2_MOUSE	P40694	mus musculu
906	35.5	21.4	373	1	GPDB_SCHPO	O14627	homo sapien	979	35.5	21.4	1026	1	BGAL_STROR	P23989	streptococ
907	35.5	21.4	380	1	IL132_HUMAN	P06623	bos taurus	980	35.5	21.4	1039	1	Y304_TREPA	O83326	treponema p
908	35.5	21.4	400	1	CNB7_BOVIN	P23466	homo sapien	981	35.5	21.4	1094	1	DPOL_GPCMV	Q69025	guinea pig
909	35.5	21.4	404	1	INBC_HUMAN			982	35.5	21.4					

983 35.5 21.4 1177 1 TAK DROME 024592 drosophila
 984 35.5 21.4 1197 1 EVGS_ECOLI P84402 escherichia
 985 35.5 21.4 1197 1 EVGS_ECOLI P8085 escherichia
 986 35.5 21.4 1316 1 RPOC_MCTU P47769 mycobacteri
 987 35.5 21.4 1612 1 RPOC_PMYV P99657 pepper mild
 988 35.5 21.4 1612 1 RPOC_PMYV P99657 pepper mild
 989 35.5 21.4 1612 1 RPOC_PMYV P99657 pepper mild
 990 35.5 21.4 1612 1 RPOC_PMYV P99657 pepper mild
 991 35.5 21.4 1612 1 RPOC_PMYV P99657 pepper mild
 992 35.5 21.4 1694 1 CLH_DICDI P12397 dicystostell
 993 35.5 21.4 2181 1 STCA_EMBI P26661 h genome po
 994 35.5 21.4 3033 1 POLG_HCVJ8 P27742 emericella
 995 35.5 21.4 3770 1 ACVS_EMBI P12955 homo sapien
 996 35.5 21.4 4377 1 ANK3_HUMAN Q14517 homo sapien
 997 35.5 21.4 4590 1 PATH_HUMAN P20767 ratu norv
 998 35 21.1 104 1 LAC2_RAT P10024 bacillus th
 999 35 21.1 114 1 YG13_BACTU Q9163 mus musculi
 1000 35 21.1 128 1 OTOR_MOUSE

ALIGNMENTS

RESULT 1
 C6ST_CHICK STANDARD: PRT: 458 AA.
 ID C6ST_CHICK
 AC Q92179: 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Chondroitin 6-sulfotransferase (EC 2.8.2.17) (C6ST).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RX STRAIN=White Leghorn; TISSUE=Embryonic chondrocytes;
 RX MEDLINE=95355490; PubMed=7629189;
 RA Fukuta M., Uchimura K., Nakashima K., Kato M., Kimata K.,
 RA Shiomura T., Habuchi O.;
 RT "Molecular cloning and expression of chick chondrocyte chondroitin 6-
 sulfotransferase";
 RL J. Biol. Chem. 270:18575-18580(1995).
 CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN
 SULFATE.
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + chondroitin =
 CC adenosine 3',5'-bisphosphate + chondroitin 6'-sulfate.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE
 CC (BY SIMILARITY).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: D49915; BAA0855.1; -
 DR Transferrase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
 FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 37 37 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 38 458 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAc...) (POTENTIAL).
 SQ SEQUENCE 458 AA: 52253 MW: 69A3B7D0A5086F0C CRC64;

Query Match 60.8%; Score 101; DB 1; Length 458;
 Best Local Similarity 59.3%; Pred. No. 3.5e-07;
 Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFEVQGLFGQHPDYFMETPAHY 29
 Db 122 RTGSFEVGEFFNOGNIFYLFEPLMH 148

RESULT 2

ARGD_METH STANDARD: PRT: 390 AA.
 ID ARGD_METH
 AC Q27392: 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetylornithine aminotransferase (EC 2.6.1.11) (ACOAT).
 GN ARGD OR MTH1337.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 NCBI_TaxID=187420;
 RX STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldridge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Viscare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-F., Rice P., Noelling J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + 2-oxoglutarate = N-
 CC acetyl-L-glutamate-5-semialdehyde + L-glutamate.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: Arginine biosynthesis; fourth step.
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AE00897; AAB8515.1; -
 DR HSSP: P12995; 10J3
 DR InterPro: IPR000954; AminoTran_3.
 DR Pfam: PF00202; aminoTran_3; 1.
 DR TIGRFAMs: TIGR00707; argD; 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; 1.
 KW Transferrase; Aminotransferase; Pyridoxal phosphate;
 KW Arginine biosynthesis; Complete proteome
 FT BINDING 246 246 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 390 AA: 41763 MW: 5F309E073DDA8FC CRC64;
 Query Match 31.9%; Score 53; DB 1; Length 390;
 Best Local Similarity 52.9%; Pred. No. 2.8;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

RESULT 3

```

NOEF RHISN
ID NOEF RHISN STANDARD: PRT: 419 AA.
AC P55472: P72326:
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Modulation protein noe (EC 2.8.2.-).
GN NOEF OR Y4HB.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234.
OC Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97361801; PubMed=9218762;
RA Hain M., Jabouri S., Quesada-Vincens S., Freilberg C., Perret X.,
RA Broughton W.J., Fellay R.,
RT Sulfation of Rhizobium sp. NGR234 nod factors is dependent on noe,
RT a new host-specificity gene.
RL Mol. Microbiol. 24:1119-1129(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freilberg C.A., Fellay R., Baitoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RT Nature 387:394-401(1997).
CC -1- FUNCTION: REQUIRED FOR THE FORMATION OF SULFATED NOD FACTOR.
CC PROPOSED TO TRANSFER ACTIVATED SULFATE (PAPS) TO THE FUCOSE
CC OF THE NOD FACTOR.
CC -1- SIMILARITY: LIMITED TO NODH AND TO C.ELEGANS FK29.8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09415; CAA70569.1; -
CC DR EMBL: AE000076; AAB91690.1; -
CC KM Modulation; Transferrase; Plasmid.
SQ SEQUENCE 419 AA; 46569 MW; 848C48B0416AAAF CRC64;

Query Match 31.6%; Score 52.5; DB 1; Length 419;
Best Local Similarity 44.0%; Pred. No. 3.5;
Matches 11; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 3 RSGSFVGOLFQHPVEYLMPEAW 27
DB 19 RSGTTLAHLQHPDITAPPEP-W 42

```

```

RESULT 4
MNTB LISIN
ID MNTB LISIN STANDARD: PRT: 240 AA.
AC C92AF9:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Manganese transport system ATP-binding protein mntB.
GN MNTB OR LIN1963.
OS Listeria innocua.
OC Bacteria: Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Vazquez-Boland J.-A., Beroche P., Bloeker H., Brandt P., Chakraborty T.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system. Probably responsible for energy coupling to the transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596170; CAC97193.1; -
CC DR Listlist; LIN01963; -
CC DR InterPro: IPR003439; ABC_transporter.
CC Pfam: PF00005; ABC_tran; 1.
CC DR ProDom: PD000006; ABC_transporter; 1.
CC DR PROSITE: PS00211; ABC_TRANSPORTER; Complete proteome.
CC KW Transport; ATP-binding; Membrane; Complete proteome.
CC NP_BIND 33 40 ATP (POTENTIAL).
SQ SEQUENCE 240 AA; 26470 MW; ADD256CE8AC20B94 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 240;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 8 FVGOLFQHPVEYLMPE 25
DB 143 FLARALQAHAFPLDEP 160

```

```

RESULT 5
MNTB LISMO
ID MNTB LISMO STANDARD: PRT: 240 AA.
AC Q8Y651:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Manganese transport system ATP-binding protein mntB.
GN MNTB OR LM01849.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Frangoul L., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";

```

```

RL Science 294:849-852(2001).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system. Probably responsible for energy coupling to the transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591981; CAC99927.1; -.
CC Listlist: LMO01849; -.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003439; ABC_transporter.
CC Pfam: PF00005; ABC_tran; 1.
CC ProDom: PD000006; ABC_transporter; 1.
CC SMART: SM00382; AAA; 1.
CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC Transprot: ATP-binding; Membrane; Complete proteome.
CC NP_BIND: 33 40 ATP (POTENTIAL).
CC SEQUENCE 240 AA; 26495 MW; C3667E78F4D853E4 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 240;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 8 FVGOLFQGHVPEYMEP 25
Db 143 FTRALAQHAETFLDEP 160
| : | : | : | : |
| : | : | : | : |

RESULT 6
GATE_MYCPN STANDARD; PRT; 338 AA.
ID GATE_MYCPN STANDARD; PRT; 338 AA.
AC P75517;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE galactose 4-epimerase).
GN GALE OR MPN257 OR MP576.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfarb R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Galactose metabolism; third step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AE000056; AAB96224.1; -.
DR HSSP: P09147; 1KVS.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
DR TIGRFAMs: TIGR01179; gale; 1.
KW Isomerase; NAD; Galactose metabolism; Complete proteome.
FT NP_BIND: 38 NAD (POTENTIAL).
SQ SEQUENCE 338 AA; 38132 MW; 9C50FF385B68C03 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 338;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 GGSFVGOLFQGHVPEY 19
Db 18 GSCFIDQLKQYDPV 32
| : | : | : | : |
| : | : | : | : |

RESULT 7
GATA_AOUAE STANDARD; PRT; 478 AA.
ID GATA_AOUAE STANDARD; PRT; 478 AA.
AC O66610;
DC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutaryl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.-) (Glu-ADT
DE subunit A).
GN GATA OR AQ_247.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
CC GIN-tRNA(GIN) THROUGH THE TRANSAMIDATION OF MISACTYLATED GLU-
CC TENA(GIN) IN ORGANISMS WHICH LACK GLUTAMINYL-tRNA SYNTHETASE. THE
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-tRNA(GIN) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Gin) + L-glutamine = ADP
CC + phosphate + L-glutamyl-tRNA(Gin) + L-glutamate.
CC -1- SUBUNIT: HETERODIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000680; AAC06569.1; -.
CC InterPro: IPR000120; Amidase.
CC InterPro: IPR004412; Gata.
CC Pfam: PF01425; Amidase; 1.
CC TIGRFAMs: TIGR00132; gata; 1.
CC PROSITE: PS00571; AMIDASES; 1.
KW Protein biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 478 AA; 53524 MW; 72E83825960E6 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 478;
Best Local Similarity 36.7%; Pred. No. 18;
Matches 11; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

```

QY 1 SWRGSSEVVGOLFQHPD-----VFLYME 24
 DB 437 AMKDGFLVGGQLIGKHWDFTLLQISYLM 466

RESULT 8

GCP2_HUMAN STANDARD: PRT: 902 AA.
 ID GCP2_HUMAN
 AC O9BSJ2; O43632;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-tubulin complex component 2 (GCP-2) (Spindle pole body protein spc97 homolog) (hspc97) (hGCP2) (h103p).
 GN GCP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=98234402; PubMed=956967;
 RA Murphy S.M., Urbani L., Stearns T.;
 RT "The mammalian gamma-tubulin complex contains homologues of the yeast spindle pole body components spc97p and spc98p.";
 RL J. Cell Biol. 141:663-674(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC Strausberg R.;
 RA TISSUE=Pancreas;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule nucleation at the centrosome.
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2, GCP3, GCP4, GCP5 and GCP6.
 CC -1- SUBCELLULAR LOCATION: Centrosome.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; may be produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; AF042379; AAC39728.1; -;
 CC EMBL; BC005011; AAH05011.1; -;
 CC KW Microtubules; Alternative splicing.
 CC FT DOMAIN 109 112
 CC FT VARSPPLIC 513 560
 CC FT LRSIKRYFLMDGDFVHMFLAEELRKPVEDIPPRLEA
 CC FT LLELAIR -> PEPSSATSSWTRATSCSTSWTSRRSSGSR
 CC FT WTSRPLAWKRSWSWRCIA (IN ISOFORM 2).
 CC FT MISSING (IN ISOFORM 2).
 CC FT VARSPPLIC 561 902
 CC FT SSEQUENCE 902 AA; 102533 MW; 4FAAF684A3758E6A CRC64;

Query Match 28.9%; Score 48; DB 1; Length 902;
 Best Local Similarity 30.6%; Pred. No. 35;
 Matches 11; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

QY 2 WRSQS-----SFVQOLFQHPDYFLMEPAWHV 29
 DB 675 WFAAGATLRLRMFLNFVONI--QYVMEVEVEPTWHI 708

RESULT 9
 GCP2_MOUSE STANDARD: PRT: 905 AA.
 ID GCP2_MOUSE
 AC O921G8;

DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-tubulin complex component 2 (GCP-2).
 GN GCP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule nucleation at the centrosome (By similarity).
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2, GCP3, GCP4, GCP5 and GCP6 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Centrosome (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; BC012519; AAH12519.1; -;
 CC DR EMBL; BC025582; AAH25582.1; -;
 CC DR MGD; MG1:1921487; 1700022B05R1K.
 CC KW Microtubules.
 CC FT SEQUENCE 905 AA; 103234 MW; E5E3F6A57EAB4357 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 905;
 Best Local Similarity 30.6%; Pred. No. 36;
 Matches 11; Conservative 8; Mismatches 7; Indels 10; Gaps 2;

QY 2 WRSQS-----SFVQOLFQHPDYFLMEPAWHV 29
 DB 675 WFAAGATLRLRMFLNFVONI--QYVMEVEVEPTWHI 708

RESULT 10
 STAN DROME STANDARD: PRT: 3579 AA.
 ID STAN DROME
 AC O9V5N8;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Prococadherin-like wing polarity protein stan precursor (Starry night protein) (Flamingo protein).
 DE STAN OR FMI OR CG1895.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Eurygata; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RX MEDLINE=200255940; PubMed=10556066;
 RA Chae J.W., Kim M.-J., Goo J.H., Collier S., Gubb D., Charlton J., Adler P.N., Park W.J.;
 RT "The prococadherin tissue polarity gene starry night encodes a member of the prococadherin family.";
 RL Development 126:5421-5429(1999).
 RN [2]
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=9418630; PubMed=10490098;
 RA Usui T., Shima Y., Shimada Y., Hirano S., Burgess R.W., Schwarz T.L.,


```

FT DOMAIN 1556 1753 LAMININ G-LIKE 1.
FT DOMAIN 1736 1792 EGF-LIKE 2 CALCIUM-BINDING.
FT DOMAIN 1796 1963 LAMININ G-LIKE 2.
FT DOMAIN 1965 2000 EGF-LIKE 3 CALCIUM-BINDING.
FT DOMAIN 2091 2126 EGF-LIKE 4 CALCIUM-BINDING.
FT DOMAIN 2744 2802 GFS.
FT DOMAIN 140 143 POLY-HIS.
FT DOMAIN 155 159 POLY-ARG.
FT DOMAIN 2567 2579 POLY-SER.
FT DOMAIN 3460 3467 POLY-GLN.
FT DISULFID 1486 1497 POTENTIAL.
FT DISULFID 1491 1506 POTENTIAL.
FT DISULFID 1508 1517 POTENTIAL.
FT DISULFID 1760 1771 POTENTIAL.
FT DISULFID 1765 1780 POTENTIAL.
FT DISULFID 1782 1791 POTENTIAL.
FT DISULFID 1969 1979 POTENTIAL.
FT DISULFID 1973 1988 POTENTIAL.
FT DISULFID 1990 1999 POTENTIAL.
FT DISULFID 2092 2095 POTENTIAL.
FT DISULFID 2097 2114 POTENTIAL.
FT DISULFID 2116 2125 POTENTIAL.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 886 886 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

Query Match 28.9%; Score 48; DB 1; Length 3579;
Best Local Similarity 42.9%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```

OY 4 GSSSFVQLEFGQHPDYIME 24

Db 3374 GSGSOTIGHMSEFHPDAATLSD 3394

```

RESULT 11
Y4IE_RHISN STANDARD; PRT; 135 AA.
AC P55488;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 15.4 kDa protein Y4IE.
GN Y4IE.
OS Rhizobium sp. (Strain NGR234).
OG Plasmid sym PNGR234.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Balroch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1 SIMILARITY: NONE OBVIOUS OUTSIDE OF SIMILARITY WITH Y4RL AND Y4ZA.
CC -1 SIMILARITY: NONE OBVIOUS OUTSIDE OF SIMILARITY WITH Y4RL AND Y4ZA.
CC POTENTIAL FRAGMENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE000078; AAB91700.1; -
CC Hypothetical protein; Plasmid.
CC SEQUENCE 135 AA; 15402 MW; 8166C909058F14C3 CRC64;
SQ

```

```

Query Match 28.3%; Score 47; DB 1; Length 135;
Best Local Similarity 43.8%; Pred. No. 6.7;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

OY 2 WRSGSFVQLEFGQHP 17

Db 76 WRTGSFSLRLQAEYP 91

RESULT 12

ID RMLB_STRMU STANDARD; PRT; 348 AA.

AC P95780;

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE GDP-glucose 4,6-dehydratase (EC 4.2.1.46).

GN RMLB.

OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1309;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-XC;

RA Tsukioaka Y., Yamashita Y., Nakano Y., Oho T., Koga T.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1 CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-

CC glucose + H(2)O.

CC -1 COFACTOR: NAD.

CC -1 PATHWAY: dTDP-L-thiamose biosynthesis.

CC -1 SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTPP-GLUCOSE

CC DEHYDRATASE SUBFAMILY.

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

CC EMBL: D78182; BAA11249.1; -

DR HSSP: P27830; 18XK.

DR InterPro: IPR001509; Epimerase_Dh.

DR Pfam: PF01370; Epimerase; 1.

DR TrEMBL: TIGR01181; dTDP-gluc_dehyd; 1.

DR Lipopolysaccharide biosynthesis; Lyase; NAD.

FT NP_BIND 11 17 NAD (POTENTIAL).

FT SEQUENCE 348 AA; 39251 MW; 034E68B8B15079DD CRC64;

Query Match 28.3%; Score 47; DB 1; Length 348;

Best Local Similarity 53.3%; Pred. No. 18;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 GSSSFVQLEFGQHPDY 19

Db 17 GSNFVHYNNHNPV 31

RESULT 13

ID CG12_CANAL

AC P43062; STANDARD; PRT; 465 AA.

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE G1/S-specific cyclin CLN2.

GN CLN2.

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI_TaxID=5476;

CC USUALLY C16, HOWEVER IN THE MAMMARY GLANDS OF NONRUMINANT MAMMALS,
 CC AND IN THE UROGYAL GLAND OF CERTAIN WATERFOWL EXISTS A SECOND
 CC THIOESTERASE WHICH RELEASE MEDIUM-CHAIN LENGTH FATTY ACIDS (C8 TO
 CC C22).
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl]-
 CC carrier protein] + oleate.
 CC -1- SIMILARITY: TO OTHER THIOESTERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ID EMBL; M12101; AAA49222.1; -;
 CC DR EMBL; M21635; AAA49219.1; -;
 CC DR PIR; A00775; ESPKTM.
 CC DR InterPro: IPR000379; Ser. esters. site.
 CC DR InterPro: IPR001031; Thioesterase.
 CC DR Pfam; PF00975; Thioesterase; 1.
 CC KW Fatty acid biosynthesis; Hydrolyase.
 CC FT ACT_SITE 90 90 PROBABLE.
 CC FT ACT_SITE 226 226 BY SIMILARITY.
 CC FT CONFLICT 43 43 V -> E (IN REF. 2).
 CC SQ SEQUENCE 251 AA; 28807 MW; 3DFA25170D9EDC50 CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 251;
 CC Best Local Similarity 47.8%; Pred. No. 15;
 CC Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
 CC
 CC QY 4 SSGSFGVGLFGQHPDVFYLMPP 26
 CC DB 214 SGGTSTSLPQNH---FYLMPPS 233
 CC
 CC RESULT 16
 CC CARD_ERMCA STANDARD; PRT; 376 AA.
 CC ID CARD_ERMCA STANDARD; PRT; 376 AA.
 CC AC O9XB58;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Carabapenem antibiotics biosynthesis protein CARD.
 CC GN CARD.
 CC OS Erwina carotovora.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Pectobacterium.
 CC OX NCBI_TaxID=354;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=ATCC 39048 / GS101;
 CC RX MEDLINE=98065591; PubMed=9402024;
 CC RA McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,
 CC Stewart G.S., Bycroft B.W., Salmon G.P.,
 CC "Analysis of the carabapenem gene cluster of Erwina carotovora:
 CC RT definition of the antibiotic biosynthetic genes and evidence for a
 CC novel beta-lactam resistance mechanism.";
 CC RL Mol. Microbiol. 26:545-556(1997).
 CC -1- PATHWAY: Carabapenem antibiotics biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE PROLINE OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ID EMBL; U17224; AAD38232.1; -;
 CC DR EMBL; U17224; AAD38232.1; -;
 CC DR InterPro: IPR002872; Pro.ch.

DR Pfam; PF01619; Pro.ch. 1.
 KW Antibiotic biosynthesis; Oxidoreductase.
 SQ SEQUENCE 376 AA; 42885 MW; 1F8AF5AE77330E7C CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 376;
 CC Best Local Similarity 36.7%; Pred. No. 23;
 CC Matches 11; Conservative 8; Mismatches 10; Indels 1; Gaps 1;
 CC
 CC QY 1 SWSGSGFVGLFGQHPDVFY-LMEPAMVY 29
 CC DB 298 SLRAGRTWGLYGLDHDITTYSLQGFVY 327
 CC
 CC RESULT 17
 CC CYMO_ACISP STANDARD; PRT; 542 AA.
 CC ID CYMO_ACISP STANDARD; PRT; 542 AA.
 CC AC P12015;
 CC DT 01-OCT-1989 (Rel. 12, Created)
 CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
 CC DT 01-AUG-1991 (Rel. 19, Last annotation update)
 CC DE Cyclohexanone monooxygenase (EC 1.14.13.22).
 CC OS Acinetobacter sp.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC OC Acinetobacter.
 CC OX NCBI_TaxID=472;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 CC RC STRAIN=NCIB 9871;
 CC RX MEDLINE=88115180; PubMed=3338974;
 CC RA Chen Y.-C.T., Peoples O.P., Walsh C.T.;
 CC "Acinetobacter cyclohexanone monooxygenase: gene cloning and sequence
 CC RT determination.";
 CC RL J. Bacteriol. 170:781-789(1988).
 CC -1- CATALYTIC ACTIVITY: Cyclohexanone + NADPH + O(2) = 6-hexanolide +
 CC NADP(+) + H(2)O.
 CC -1- COFACTOR: FAD.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ID EMBL; M19029; AAA21892.1; -;
 CC DR PIR; A28550; A28550.
 CC DR InterPro: IPR000960; Flav.conc._mnoxgn.
 CC DR Pfam; PF00743; FMO-like; 1.
 CC KW Aromatic hydrocarbons catabolism; Oxidoreductase; Monooxygenase;
 CC KW NADP; Flavoprotein; FAD.
 CC FT INIT_MET 0 0
 CC FT IN_BIND 6 18 FAD (BY SIMILARITY).
 CC FT NP_BIND 176 208 NADP (BY SIMILARITY).
 CC SQ SEQUENCE 542 AA; 60760 MW; 7ACDD25C7BD388F CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 542;
 CC Best Local Similarity 32.0%; Pred. No. 34;
 CC Matches 8; Conservative 8; Mismatches 8; Indels 1; Gaps 1;
 CC
 CC QY 2 WRSQ-SFVGLFGQHPDVFYLMPP 25
 CC DB 403 WKBPSSYMGVYNNYNNMFMVILGP 427
 CC
 CC RESULT 18
 CC ADCC_STRPN STANDARD; PRT; 234 AA.
 CC ID ADCC_STRPN STANDARD; PRT; 234 AA.
 CC AC O87862;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Zinc transport system ATP-binding protein adcc.

GN ADCC OR SP2171.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus
 NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R6;
 RX MEDLINE=98438930; PubMed=9765793;
 RA Dintilhac A., Claverys J.-P.;
 RT "The *adc* locus, which affects competence for genetic transformation in
 RT Streptococcus pneumoniae, encodes an ABC transporter with a putative
 RT lipoprotein homologous to a family of streptococcal adhesins.";
 RL Res. Microbiol. 148:119-131(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Hatt D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Hickey E.K.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=98025470; PubMed=9379902;
 RA Dintilhac A., Allouin G., Granadel C., Claverys J.-P.;
 RT "Competence and virulence of Streptococcus pneumoniae: *Adc* and *Psaa*
 RT mutants exhibit a requirement for Zn and Mn resulting from
 RT inactivation of putative ABC metal permeases.";
 RL Mol. Microbiol. 25:727-739(1997).
 CC -1- FUNCTION: PART OF THE ATP-DRIVEN TRANSPORT SYSTEM ADABC FOR ZINC.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -1-
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z11552; CA96186.1; -
 DR EMBL: AE007505; AA76225.1; -
 DR TIGR: SP2171; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran.1.
 DR ProDom: PD000006; ABC_transportr; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR ATP-binding; transport; zinc transport; Complete proteome.
 SQ SEQUENCE 234 AA; 26542 MW; 2AF782EP93D5C1A0 CRC64;

Query Match 27.7%; Score 46; DB 1; Length 234;
 Best Local Similarity 47.1%; Pred. No. 17;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 9 VGOILFGHPDYVYLMEMP 25
 DB 148 IARFASDPDYVILDEP 164

RESULT 19
 CYAB-LEIDO
 ID CYAB-LEIDO
 AC Q25263; STANDARD; PRT; 1331 AA.

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Receptor-type adenylyl cyclase B (EC 4.6.1.1) (ATP pyrophosphate-
 DE lyase) (Adenylyl cyclase).
 GN RAC-B.
 OS Leishmania donovani.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5661;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IS Sudanese;
 RX MEDLINE=95340554; PubMed=7615561;
 RA Sanchez M.A., Zeoli D., Klamo E.M., Kavanaugh M.P., Landfear S.M.;
 RT "A family of putative receptor-adenylyl cyclases from Leishmania
 RT donovani.";
 RL J. Biol. Chem. 270:17551-17558(1995).
 CC -1- FUNCTION: COULD ACT AS A RECEPTOR FOR A UNKNOWN LIGAND.
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE INSECT STAGE (PROMASTIGOTE)
 CC BUT NOT IN THE MAMMALIAN HOST STAGE OF THE PARASITE LIFE CYCLE.
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.
 CC -1-
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U17043; AA74999.1; -
 DR InterPro: IPR001054; G_cyclase.
 DR Pfam: PF00211; guanylate_cyc; 1.
 DR SMART: SM00044; CYCC; 1.
 DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
 KW lysase; cAMP synthesis; Transmembrane; Receptor; Glycoprotein.
 FT DOMAIN 1 33
 FT TRANSMEM 34 33
 FT DOMAIN 34 898
 FT TRANSMEM 899 919
 FT DOVAIN 920 1331
 FT CARBOHYD 255 255
 FT CARBOHYD 429 429
 FT CARBOHYD 558 558
 FT CARBOHYD 574 574
 FT CARBOHYD 657 657
 SQ SEQUENCE 1331 AA; 144162 MW; 5FC3AA22DC1E5072 CRC64;

Query Match 27.7%; Score 46; DB 1; Length 1331;
 Best Local Similarity 36.4%; Pred. No. 1e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 6 SSEVGOILFGHPDYVYLMEMP 27
 DB 390 AAYVNNFNEHPCVTQMLMGW 411

RESULT 20
 AATM-CHICK
 ID AATM-CHICK
 AC P00508; STANDARD; PRT; 423 AA.
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
 DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 NCBI_TaxID=9031;

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=86059504; PubMed=3840803;
RA Jausasi R., Cotton B., Juretic N., Christen P., Schumperli D.;
RT "The primary structure of the precursor of chicken mitochondrial
RL J. Biol. Chem. 260:16060-16063(1985).
RN [2]
RX MEDLINE=89238520; PubMed=6345546;
RA Graf-Hausner U., Wilson K.J., Christen P.;
RT "The covalent structure of mitochondrial aspartate aminotransferase
RL from chicken. Identification of segments of the polypeptide chain
RN invariant specifically in the mitochondrial isoenzyme.";
RX J. Biol. Chem. 258:8813-8826(1983).
RN [3]
RP GENE STRUCTURE.
RC STRAIN=White leghorn;
RX MEDLINE=90382432; PubMed=2401287;
RA Juretic N., Mattes U., Zlak M., Christen P., Jausasi R.;
RT "Structure of the genes of two homologous intracellularly heterotopic
RL isoenzymes. Cytosolic and mitochondrial aspartate aminotransferase of
RN chicken.";
RL Eur. J. Biochem. 192:119-126(1990).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=92277655; PubMed=1593633;
RA McPhalen C.A., Vincent M.G., Jansonius J.N.;
RT "X-ray structure refinement and comparison of three forms of
RL mitochondrial aspartate aminotransferase.";
RN J. Mol. Biol. 225:495-517(1992).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxalacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation in
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M12105; AAA48603.1; -.
DR PIR: A00597; XNCHDM.
DR PIR: A24554; A24554.
DR PIR: S11346; S11346.
DR PDB: 7AAT; 3I-OCT-93.
DR PDB: 8AAT; 3I-OCT-93.
DR PDB: 9AAT; 3I-OCT-93.
DR PDB: 1AKA; 3I-JUL-94.
DR PDB: 1AKB; 3I-JUL-94.
DR PDB: 1AKC; 3I-JUL-94.
DR PDB: 1AMA; 3I-OCT-93.
DR PDB: 1MAP; 3I-JAN-94.
DR PDB: 1MAQ; 3I-JAN-94.
DR PDB: 1TAR; 3I-JAN-94.
DR PDB: 1TAS; 3I-JAN-94.
DR PDB: 1TAT; 3I-JAN-94.
DR PDB: 1OXO; 10-JUN-96.
DR PDB: 1OXP; 10-JUN-96.
DR PDB: 1IVR; 23-JUL-97.
DR InterPro: IPR004839; AminoTransf1/2.
DR InterPro: IPR000796; AsptTransf_sub.
DR InterPro: IPR004838; NHTransf_1.
DR Pfam: PF00155; aminotran_1.2; 1.
DR PRINTS: PRO0799; TRANSAMINASE.

```

DR	PROSITE	PS00105; AA:TRANSFER_CLASS.1; 1.	Transferrase; Aminoctransferase; Pyridoxal phosphate; Mitochondrion;
KW	Transit peptide; 3D-structure.		
FT	TRANSIT	1 22	MITOCHONDRION.
FT	CHAIN	23 423	ASPARTATE AMINOTRANSFERASE.
FT	BINDING	272 272	PYRIDOXAL PHOSPHATE.
FT	CONFLICT	67 67	S -> P (IN REF. 2).
FT	CONFLICT	168 168	Q -> E (IN REF. 2).
FT	CONFLICT	216 216	Q -> E (IN REF. 2).
FT	TURN	27 28	
FT	TURN	36 37	
FT	HELIIX	38 46	
FT	TURN	50 51	
FT	STRAND	53 54	
FT	STRAND	61 62	
FT	TURN	63 65	
FT	STRAND	66 67	
FT	HELIIX	71 83	
FT	TURN	84 84	
FT	TURN	92 93	
FT	HELIIX	96 107	
FT	HELIIX	112 116	
FT	TURN	117 117	
FT	STRAND	119 125	
FT	HELIIX	126 141	
FT	STRAND	147 152	
FT	TURN	156 157	
FT	HELIIX	158 164	
FT	TURN	165 166	
FT	STRAND	168 173	
FT	STRAND	175 176	
FT	TURN	177 180	
FT	STRAND	181 182	
FT	HELIIX	184 192	
FT	TURN	193 193	
FT	TURN	196 197	
FT	STRAND	199 203	
FT	TURN	208 210	
FT	HELIIX	216 229	
FT	TURN	230 230	
FT	STRAND	232 237	
FT	TURN	240 241	
FT	HELIIX	247 250	
FT	TURN	251 251	
FT	HELIIX	252 259	
FT	TURN	260 261	
FT	STRAND	265 269	
FT	TURN	271 274	
FT	HELIIX	277 279	
FT	STRAND	281 287	
FT	HELIIX	291 309	
FT	HELIIX	315 324	
FT	TURN	325 325	
FT	HELIIX	327 358	
FT	TURN	359 359	
FT	HELIIX	365 369	
FT	STRAND	374 376	
FT	HELIIX	381 391	
FT	STRAND	393 394	
FT	STRAND	396 396	
FT	TURN	397 399	
FT	STRAND	400 402	
FT	HELIIX	403 405	
FT	TURN	408 410	
FT	HELIIX	411 422	
SO	SEQUENCE	423 AA: 47241 MW: 59D65D4ED4DDF8BA CRC64;	
Query Match	27.4%;	Score 45.5; DB 1; Length 423;	
Best Local Similarity	40.7%;	Pred. No. 37;	
Matches 11;	Conservative 6;	Mismatches 9; Indels 1; Gaps 1	
1	SMRSGSFEVGLFGQHDPVFTIMEPAW	27	
1	:	:	

DB 130 SLRVCANFLQRFKFSRDV-YLPKPSM 155

RESULT 21

AATM_MOUSE

ID AATM_MOUSE STANDARD: PRT: 430 AA.

AC P05202; 009188; 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)

DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).

GN GOT2 OR GOT-2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87057413; PubMed=3782150;

RX Obaru K., Nomiyama H., Shimada K., Nagashima F., Morino Y.;

RT "Cloning and sequence analysis of mRNA for mouse aspartate aminotransferase isoenzymes."

RT J. Biol. Chem. 261:16976-16983(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/He; TISSUE=Liver;

RX MEDLINE=88118911; PubMed=2828632;

RA Tsuzuki T., Obaru K., Setoyama C., Shimada K.;

RT "Structural organization of the mouse mitochondrial aspartate aminotransferase gene."

RT J. Mol. Biol. 198:21-31(1987).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB; TISSUE=Liver;

RA Bradbury M.W., Berk P.D.;

RT Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate + L-glutamate.

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A MITOCHONDRIAL ONE.

CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J02622; AAA37264.1; -

DR EMBL: X06917; CAA30015.1; -

DR EMBL: X06918; CAA30015.1; JOINED.

DR EMBL: X06919; CAA30015.1; JOINED.

DR EMBL: X06920; CAA30015.1; JOINED.

DR EMBL: X06921; CAA30015.1; JOINED.

DR EMBL: X06922; CAA30015.1; JOINED.

DR EMBL: X06923; CAA30015.1; JOINED.

DR EMBL: X06924; CAA30015.1; JOINED.

DR EMBL: X06925; CAA30015.1; JOINED.

DR EMBL: X06926; CAA30015.1; JOINED.

DR EMBL: M37259; AAA37265.1; ALT-SEQ.

DR EMBL: M37250; AAA37265.1; JOINED.

DR EMBL: M37251; AAA37265.1; JOINED.

DR EMBL: M37252; AAA37265.1; JOINED.

DR EMBL: M37253; AAA37265.1; JOINED.

DR EMBL: M37254; AAA37265.1; JOINED.

DR EMBL: M37255; AAA37265.1; JOINED.

DR EMBL: M37256; AAA37265.1; JOINED.

DR EMBL: M37258; AAA37265.1; JOINED.

DR EMBL: U82470; AAB91426.1; -

DR PIR: S01174; S01174.

DR PIR: A25349; A25349.

DR HSSP: P00508; TAAT.

DR MGD: MGI:95792; GOT2.

DR InterPro: IPR004839; AminoTransf1/2.

DR InterPro: IPR000796; Aspartate sub.

DR InterPro: IPR004838; NHTransf1.

DR Pfam: PF00155; AminoTransf1_2; 1.

DR PRINTS: PRO0199; TRANSAMINASE.

DR PROSITE: PS00105; AA-TRANSFER-CLASS-1; 1.

RK TRANSFERASE; Aminotransferase; Pyridoxal phosphate; Mitochondrion;

KW Trans peptide.

FT TRANSIT 1 29 MITOCHONDRION.

FT CHAIN 30 430 ASPARTATE AMINOTRANSFERASE.

FT BINDING 279 279 PYRIDOXAL PHOSPHATE (BY SIMILARITY).

FT CONFLICT 146 146 Q -> E (IN REF. 3).

SO SEQUENCE 430 AA; 47411 MW; D590524CA/FFB885 CRC64;

Query Match 27.4%; Score 45.5; DB 1; Length 430;

Best Local Similarity 44.0%; Pred. No. 38;

Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

OY 3 RSGSFTVGQLFGHPDFTLMEPM 27

DB 139 RVGASFLORFKFSRDV-LPKPSM 162

RESULT 22

AATM_RAT

ID AATM_RAT STANDARD: PRT: 430 AA.

AC P00507; Q64551; 21-JUL-1986 (Rel. 01, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)

DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).

GN GOT2 OR MAAT.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88106546; PubMed=3322287;

RA Mattingly J.R. Jr., Rodriguez-Betrocot F.J., Gordon J., Iriarte A., Martinez-Carrion M.;

RT "Molecular cloning and in vivo expression of a precursor to rat mitochondrial aspartate aminotransferase."

RT Biochem. Biophys. Res. Commun. 149:859-865(1987).

RN [2]

RP SEQUENCE OF 30-430.

RX MEDLINE=81133608; PubMed=7470110;

RA Huynh Q.K., Sakakibara R., Watanabe T., Wada H.;

RT "Primary structure of mitochondrial glutamic oxaloacetic transaminase from rat liver: comparison with that of the pig heart isozyme.";

RT Biochem. Biophys. Res. Commun. 97:474-479(1980).

RN [3]

RP SEQUENCE OF 30-430.

RX MEDLINE=82075716; PubMed=7309704;

RA Huynh Q.K., Sakakibara R., Watanabe T., Wada H.;

RT "The complete amino acid sequence of mitochondrial glutamic oxaloacetic transaminase from rat liver.";

RT J. Biochem. 90:863-875(1981).

RN [4]

RP SEQUENCE OF 1-30 FROM N.A.

RC STRAIN=Wistar; TISSUE=Liver;

RX MEDLINE=95279397; PubMed=7759512;

RA Jiang H.H., Costello L.C., Franklin R.B.;

RT "Androgen modulation of multiple transcription start sites of the mitochondrial aspartate aminotransferase gene in rat prostate.";

DE RUBISCO operon transcriptional regulator.
 GN CBR.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSM419;
 RA Penner B.J., Tiwari R.P., Dilworth M.J.;
 RT Genetic regulation of C₂ metabolism in Sinorhizobium meliloti.
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR FOR THE CBB OPERON FOR
 CC RUBISCO AND OTHER CALVIN CYCLE GENES.
 CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF21846; AAF25374.1; -
 CC InterPro: IPR000847; HTH_LysR.
 CC InterPro: IPR005119; LysR_subst.
 CC Pfam: PF00126; HTH.1; 1.
 CC DR Pfam: PF03466; LysR_substrate; 1.
 CC DR PRINTS: PRO0039; HTHLYSR.
 CC DR PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
 CC KW Transcription regulation; Activator; DNA-binding.
 CC FT DNA_BIND 21 40 H-T-H MOTIF (POTENTIAL).
 CC SQ SEQUENCE 313 AA; 34192 MW; 818CF5EC0FCED24 CRC64;
 OY Query Match 27.1%; Score 45; DB 1; Length 313;
 Db Best Local Similarity 55.6%; Pred. No. 32;
 Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;
 OY 11 QLFQGDVPEVLYMEPAWH 28
 Db 160 QVFQDHLVF--IAPAGH 175
 RESULT 27
 AROC_STYX3
 ID AROC_STYX3 STANDARD; PRT; 362 AA.
 AC P23353;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chorismate synthase (EC 4.6.1.4) (5-enolpyruvylshikimate-3-phosphate
 DE phospholase).
 GN AROC OR SL11747.
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94086566; PubMed=7505271;
 RA Schmidt J., Bubunenko M., Subramanian A.R.;
 RT "A novel operon organization involving the genes for chorismate
 RT synthase (aromatic biosynthesis pathway) and ribosomal GTPase center
 RT proteins (L11, L1, L10, L12: rplKarl) in cyanobacterium Synecocystis
 RT PCC 6803".
 RL J Biol. Chem. 268:27447-27457(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Keneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugijara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Natsu K.,

RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions".
 RL DNA Res. 3:109-136(1996).
 RN [3]
 RP SEQUENCE OF 1-138 FROM N.A.
 RX MEDLINE=91002677; PubMed=2119815;
 RA Sibold C., Subramanian A.R.;
 RT Cloning and characterization of the genes for ribosomal proteins L10
 RT and L12 from Synecocystis sp. PCC 6803: comparison of gene
 RT clustering pattern and protein sequence homology between
 RT cyanobacteria and chloroplasts."
 RL Biochim. Biophys. Acta 1050:61-68(1990).
 CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -1- COFACTOR: REDUCED FLAVIN (BY SIMILARITY).
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
 CC seventh step.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67516; GAA47855.1; -
 CC EMBL: D90906; BAA17415.1; -
 CC EMBL: X53178; CAA37319.1; -
 CC PIR: S13070; S13070.
 CC InterPro: IPR000453; Chorismate-synt.
 CC DR Pfam: PF01264; Chorismate-synt; 1.
 CC DR ProDom: PD002941; Chorismate-synt; 1.
 CC DR TIGRPFAMs: TIGR000033; aroc; 1.
 CC DR PROSITE: PS00787; CHORISMATE_SYNTHASE_1; 1.
 CC DR PROSITE: PS00788; CHORISMATE_SYNTHASE_2; 1.
 CC DR PROSITE: PS00789; CHORISMATE_SYNTHASE_3; 1.
 CC KW Lyase; Aromatic amino acid biosynthesis; Complete proteome.
 CC SQ SEQUENCE 362 AA; 39287 MW; 9709B0F409168AB8 CRC64;
 OY Query Match 27.1%; Score 45; DB 1; Length 362;
 Db Best Local Similarity 50.0%; Pred. No. 37;
 Matches 11; Conservative 1; Mismatches 8; Indels 2; Gaps 1;
 OY 5 GSGFVGOLF--GOHPVPEYIME 24
 Db 249 GSGFAGTLLTGSQHNDEYILDE 270
 RESULT 28
 OPS1_SCHGR
 ID OPS1_SCHGR STANDARD; PRT; 381 AA.
 AC Q94741;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Opsin 1.
 GN IOL.
 OS Schistocerca gregaria (Desert locust).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pserygota; Neoptera; Orthoptera; Orthoptera; Caelifera;
 OC Acridoidea; Acrididae; Acrididae; Cyrtacanthacridinae;
 OC Schistocerca.
 OX NCBI_TaxID=7010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97301174; PubMed=9156194;
 RA Townier P., Harris P., Wolstenholme A.J., Hill C., Worm K., Gartner W.,

RT "Primary structure of locust opsins: a speculative model which may
 account for ultraviolet wavelength light detection.";
 RL Vision Res. 37:495-503(1997).
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
 MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
 LINKED TO CIS-RETINAL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
 BE PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC OPSIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X80071; CAA56377.1; -
 DR HSSP: P02699; 1F88.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL2; 1.
 DR PROSITE: PS00238; OPSIN; 1.
 KW Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;
 KW Glycoprotein; G-protein coupled receptor; Vision.
 FT DOMAIN 1 53
 FT TRANSMEM 54 78
 FT DOMAIN 79 90
 FT TRANSMEM 91 115
 FT DOMAIN 116 130
 FT TRANSMEM 131 150
 FT DOMAIN 151 169
 FT TRANSMEM 170 193
 FT DOMAIN 194 217
 FT TRANSMEM 218 245
 FT DOMAIN 246 280
 FT TRANSMEM 281 304
 FT DOMAIN 305 311
 FT TRANSMEM 312 336
 FT DOMAIN 337 381
 FT DISULFID 127 204
 FT BINDING 323 323
 FT CARBOHYD 24 24
 FT CARBOHYD 200 200
 SQ SEQUENCE 381 AA; 42398 MW; 50D38FF9038D5BA3 CRC64;
 Query Match 27.1%; Score 45; DB 1; Length 381;
 Best Local Similarity 27.6%; Pred. No. 39;
 Matches 8; Conservative 5; Mismatches 14; Indels 2; Gaps 1;
 QY 2 WRSGSFVGLFGQH--PDVFILMEPAWH 28
 DB 16 WGGSGGFANQTVVKKVPEMLYLPDHW 44
 RESULT 29
 FIBA_HUMAN STANDARD: PRT; 866 AA.
 AC P02671; Q9BX62; Q9UCH2;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen alpha/alpha-E chain precursor [contains: Fibrinopeptide A].
 GN FGA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ALPHA-E FORM).
 RX MEDLINE-93090725; PubMed-1457396;
 RA Fu Y., Weissbach L., Plant P.W., Oddoux C., Cao Y., Liang T.J.,
 RA Roy S.N., Reeman C.M., Grieninger G.,
 RT "Carboxy-terminal extended variant of the human fibrinogen alpha
 subunit: a novel exon conferring marked homology to beta and gamma
 subunits.";
 RT Biochemistry 31:11968-11972(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ALPHA-E FORM).
 RA Chung D.W., Grieninger G.;
 RT "Fibrinogen DNA and protein sequences.";
 RL (in) Edert R.F. (eds.);
 RL Index of variant human fibrinogens, pp.13-24, CRC Press,
 RL Boca Raton (1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ALPHA-E; ALPHA), AND VARIANTS VAL-6; ALA-331 AND
 RP ALA-456.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-655 FROM N.A. (ALPHA-E FORM).
 RX TISSUE=Liver;
 RT MEDLINE-91344740; PubMed-2102623;
 RA Chung D.W., Harris J.E., Davie E.W.;
 RT "Nucleotide sequences of the three genes coding for human
 fibrinogen.";
 RT Adv. Exp. Med. Biol. 281:39-48(1990).
 RN [5]
 RP SEQUENCE FROM N.A. (ALPHA FORM).
 RX MEDLINE-93247396; PubMed-6575389;
 RA Kant J.A., Lord S.T., Crabtree G.R.;
 RT "Partial mRNA sequences for human A alpha, B beta, and gamma
 fibrinogen chains: evolutionary and functional implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3953-3957(1983).
 RN [6]
 RP SEQUENCE OF 1-629 FROM N.A.
 RX MEDLINE-93268432; PubMed-6688355;
 RA Rixon M.W., Chan W.-Y., Davie E.W., Chung D.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 the alpha chain of human fibrinogen.";
 RL Biochemistry 22:3237-3244(1983).
 RN [7]
 RP SEQUENCE OF 20-629.
 RA Herschen A., Iofftspeich F., Southan C., Topfer-Petersen E.;
 RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some
 structural variants.";
 RL (in) Peeters H. (eds.);
 RL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,
 RL Pergamon Press, Oxford (1980).
 RN [8]
 RP SEQUENCE OF 20-629, AND DISULFIDE BONDS.
 RX MEDLINE-90088231; PubMed-518846;
 RA Walt K.W.K., Cottrell B.A., Strong D.D., Doonittle R.F.;
 RT "Amino acid sequence studies on the alpha chain of human
 fibrinogen: overlapping sequences providing the complete sequence.";
 RL Biochemistry 18:5410-5416(1979).
 RN [9]
 RP SEQUENCE OF 110-156 FROM N.A.
 RX MEDLINE-84065777; PubMed-6689067;
 RA Imam A.M., Eaton M.A., Williamson R., Humphries S.;
 RT "Isolation and characterization of cDNA clones for the A alpha- and
 gamma-chains of human fibrinogen.";
 RL Nucleic Acids Res. 11:7427-7434(1983).
 RN [10]
 RP SEQUENCE OF 605-644 FROM N.A. (ALPHA FORM).
 RX MEDLINE-93254384; PubMed-6575700;
 RA Chung D.W., Rixon M.W., Que B.G., Davie E.W.;
 RT "Cloning of fibrinogen genes and their cDNA.";
 RL Ann. N.Y. Acad. Sci. 408:449-456(1983).

[11] SEQUENCE OF 20-35.
RA Blomback B., Blomback M., Grondahl N.U., Guthrie C., Hinton M.;
RT "Studies on fibrinopeptides from primates.";
RL Acta Chem. Scand. 19:1788-1789(1965).
RN [12]
RP CROSS-LINKING ACCEPTOR SITES.
RX MEDLINE=80088230; PubMed=518845;
RA Cottrell B.A., Strong D.D., Watt K.W.K., Doolittle R.F.;
RT "Amino acid sequence studies on the alpha chain of human fibrinogen.
RT Exact location of cross-linking acceptor sites.";
RL Biochemistry 18:5405-5410(1979).
RN [13]
RP CROSS-LINKING ACCEPTOR SITES.
RX MEDLINE=78130085; PubMed=632262;
RA Fretto L.J., Ferguson E.W., Steinman H.M., McKee P.A.;
RT "Localization of the alpha-chain cross-link acceptor sites of human
RT fibrin.";
RL J. Biol. Chem. 253:2184-2195(1978).
RN [14]
RP VARIANT, AND DISULFIDE BONDS.
RX MEDLINE=76225080; PubMed=936108;
RA Blomback B., Hessel B., Hogg D.;
RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
RL Thromb. Res. 8:639-658(1976).
RN [15]
RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
RX MEDLINE=84305751; PubMed=6383194;
RA Doolittle R.F.;
RT "Fibrinogen and fibrin.";
RL Annu. Rev. Biochem. 53:195-229(1984).
RN [16]
RP CROSS-LINKING SITE FOR ALPHA-2-PLASMIN INHIBITOR.
RX MEDLINE=87057190; PubMed=2877981;
RA Kimura S., Aoki N.;
RT "Cross-linking site in fibrinogen for alpha 2-plasmin inhibitor.";
RL J. Biol. Chem. 261:15591-15595(1986).
RN [17]
RP PHOSPHORYLATION.
RX MEDLINE=84104274; PubMed=6318767;
RA Itarte E., Plana M., Guasch M.D., Martos C.;
RT "Phosphorylation of fibrinogen by casein kinase 1.";
RL Biochem. Biophys. Res. Commun. 117:631-636(1983).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-39.
RX MEDLINE=92218459; PubMed=1560020;
RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
RT "The structure of residues 7-16 of the A alpha-chain of human
RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";
RL J. Biol. Chem. 267:7911-7920(1992).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 130-216.
RX MEDLINE=97472408; PubMed=9333233;
RA Spraggon G., Everse S.J., Doolittle R.F.;
RT "Crystal structures of fragment D from human fibrinogen and its
RT crosslinked counterpart from fibrin.";
RL Nature 389:455-462(1997).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 130-216.
RX MEDLINE=98292395; PubMed=9628725;
RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
RT "Crystal structure of fragment double-D from human fibrin with two
RT different bound ligands.";
RL Biochemistry 37:8637-8642(1998).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 670-866.
RX MEDLINE=98356117; PubMed=9689040;
RA Spraggon G., Applegate D., Everse S.J., Zhang J.Z., Veerapandian L.,
RA Redman C., Doolittle R.F., Griening G.;
RT "Crystal structure of a recombinant alpha2C domain from human
RT fibrinogen-420.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:9099-9104(1998).
RN [22]

RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=99175089; PubMed=10074346;
RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
RT "Conformational changes in fragments D and double-D from human
RT fibrin(ogen) upon binding the peptide ligand Gly-Arg-Pro-amide.";
RL Biochemistry 38:2941-2946(1999).
RN [23]
RP VARIANT KYOTO-2.
RX MEDLINE=91300048; PubMed=2070049;
RA Yoshida N., Okuma M., Hirata H., Matsuda M., Yamazumi K., Asakura S.;
RT "Fibrinogen Kyoto II, a new congenitally abnormal molecule,
RT characterized by the replacement of A alpha proline-18 by leucine.";
RL Blood 78:149-153(1991).
RN [24]
RP VARIANT LIMA.
RX MEDLINE=92340680; PubMed=1634621;
RA Maekawa H., Yamazumi K., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., Arocha-Pinango C.L., Rodriguez S., Nagy H.,
RA Perez-Requejo J.L., Matsuda M.;
RT "Fibrinogen Lima: a homozygous dysfibrinogen with an A
RT alpha-arginine-141 to serine substitution associated with extra
RT N-glycosylation at A alpha-asparagine-139. Impaired fibrin gel
RT formation but normal fibrin-facilitated plasminogen activation
RT catalyzed by tissue-type plasminogen activator.";
RL J. Clin. Invest. 90:67-76(1992).
RN [25]
RP VARIANT CARACAS-2.
RX MEDLINE=91268018; PubMed=1675636;
RA Maekawa H., Yamazumi K., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., de Bosch N.B., Carvajal Z., Ojeda A.,
RA Arocha-Pinango C.L., Matsuda M.;
RT "An A alpha Ser-434 to N-glycosylated Asn substitution in a
RT dysfibrinogen, fibrinogen Caracas II, characterized by impaired
RT fibrin gel formation.";
RL J. Biol. Chem. 266:11575-11581(1991).
RN [26]
RP VARIANT DUSART.
RX MEDLINE=93232289; PubMed=8473507;
RA Koopman J., Haverkate F., Grimbergen J., Lord S.T., Mosseson M.W.,
RA Diorio J.P., Stebenlist K.S., Legrand C., Soria J., Soria C.,
RA Caen J.P.;
RT "Molecular basis for fibrinogen Dusart (A alpha 554 Arg-->Cys) and
RT its association with abnormal fibrin polymerization and
RT thrombophilia.";
RN [27]
Query Match 27.1%; Score 45; DB 1; Length 866;
Best Local Similarity 52.9%; Pred. No. 93;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 SWRSGSFVGLGQHP 17
DB 333 SWRSGSGTGTSTGNQNP 349
RESULT 30
CHRD_BRARE
ID CHRD_BRARE STANDARD; PRT; 940 AA.
AC 057472; OSDEDB;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chordin precursor (Chordin protein).
GN CHD.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7935;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC TISSUE-Gastrula;
RX MEDLINE=98104254; PubMed=9441687;

```

RA Miller-Bertoglio V.E., Fisher S., Sanchez A., Mullins M.C.,
RA Halpern M.E.,
RT "Differential regulation of chordin expression domains in mutant
RT zebrafish."
RT Dev. Biol. 192:537-550(1997).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RA Fujii R., Hibi M., Hirano T., Shimizu T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Dorsalizing factor. Key developmental protein that
CC dorsalizes early vertebrate embryonic tissues by binding to
CC ventralizing TGF-beta family bone morphogenetic proteins (BMPs)
CC and sequestering them in latent complexes (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: During gastrulation, levels are highest in
CC the organizer region. Also present in the developing brain and in
CC paraxial mesoderm and ectoderm.
CC -1- DEVELOPMENTAL STAGE: First detected shortly after the midblastula
CC transition. Levels increase during gastrulation, persist through
CC early somitogenesis, but then decrease and are gone by 24
CC hours.
CC -1- SIMILARITY: BELONGS TO THE CHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 4 WFGC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL; AF034606; AAB93465.1; -
DR EMBL; AB043968; BAB1642.1; -
DR ZFIN; ZDB-GENE-990415-33; chd.
DR InterPro: IPR001007; WFG_C.
DR Pfam; PF00093; WFG; 4.
DR SMART; SM00214; WMC; 4.
DR PROSITE; PS01208; WFGC; 3.
KW Developmental protein; Repeat; Glycoprotein; Signal.
FT FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 940 CHORDIN.
FT DOMAIN 42 118 WFGC 1.
FT DOMAIN 689 748 WFGC 2.
FT DOMAIN 767 836 WFGC 3.
FT DOMAIN 855 919 WFGC 4.
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 940 AA; 104999 MW; B855CAPB4F623AC CRC64;

Query Match 27.1% Score 45; DB 1; Length 940;
Best Local Similarity 40.0% Pred. No. 1e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 8; Gaps 1;

y 4 SSSFFVQGLFGQHPDVFYLPMPAWH 28
y |||||: ||:| ||
Db 42 SGCSTFGGR-----FYSLEDTHW 58

RESULT 31
PRPD_HUMAN
AC P23468; STANDARD; PRT; 1912 AA.
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase delta precursor (EC 3.1.3.48) (R-PTP-
DE delta).
GN PRPD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_taxid=9606;

```

```

[1]
RN SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.
RX MEDLINE=95204468; Pubmed=7896816;
RA Pulido R., Krueger N.X., Seris-Pages C., Saito H., Streuli M.;
RT "Molecular characterization of the human transmembrane N-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms.";
RL J. Biol. Chem. 270:6722-6728(1995).
(2)
RN SEQUENCE OF 390-1912 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91006018; Pubmed=2170109;
RA Krueger N.X., Streuli M., Saito H.;
RT "Structural diversity and evolution of human receptor-like protein tyrosine phosphatases.";
RL EMO J. 9:3241-3252(1990).
CC -1 CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein tyrosine + phosphate.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 ALTERNATIVE PRODUCTS: A number of isoforms are produced by alternative splicing.
CC -1 Ptm: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN FROM THE TRANSMEMBRANE SEGMENT.
CC -1 SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1 SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1 SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.ch).
CC -----
DR EMBL, L38929; AAC41749.1; -.
DR EMBL, X54133; CA38068.1; -.
DR PIR, S12052; S12052.
DR HSSP, P18052; IYFO.
DR Genew: HGNC:9668; PTPRD.
DR MIM: 601598; -.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_repeat.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003598; IG_C2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_Pp.
DR Pfam: PF00041; fn3; 8.
DR Pfam: PF00047; Iy; 3.
DR Pfam: PF00102; Y_phosphatase; 2.
DR PRINTS: PR00014; FNTPeIII.
DR PRINTS: PR00700; PRTPpHPHTASE.
DR SMART: SMO0060; FN3; 8.
DR SMART: SMO0408; IGC2; 3.
DR SMART: SMO0194; PTPC; 2.
DR -PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
DR -PROSITE: PS50056; TYR_PHOSPHATASE_2; 2.
DR -PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 2.
KW Hydroxylase; Receptor; Glycoprotein; Signal; Transmembrane; Repeat; Immunoglobulin domain; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 1912
FT DOMAIN 21 1265
FT TRANSSEM 1266 1290
FT DOMAIN 1291 1912
FT DOMAIN 23 115
FT DOMAIN 118 225
FT DOMAIN 232 318
FT DOMAIN 320 414
FT DOMAIN 417 513
FT DOMAIN 516 606
FT DOMAIN 609 708
FT DOMAIN

```

```

FT DOMAIN 711 822 FIBRONECTIN TYPE-III 5.
FT DOMAIN 825 916 FIBRONECTIN TYPE-III 6.
FT DOMAIN 918 1017 FIBRONECTIN TYPE-III 7.
FT DOMAIN 1020 1137 FIBRONECTIN TYPE-III 8.
FT DOMAIN 1375 1618 PROTEIN-TYROSINE PHOSPHATASE 1.
FT DOMAIN 1619 1912 PROTEIN-TYROSINE PHOSPHATASE 2.
FT ACT_SITE 1553 1553 BY SIMILARITY.
FT ACT_SITE 1844 1844 BY SIMILARITY.
FT SITE 1175 1178 CLEAVAGE (POTENTIAL).
FT CARBOXYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOXYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOXYD 724 724 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOXYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASAPLIC 181 189 MISSING (IN KIDNEY ISOFORM).
FT VASAPLIC 229 229 MISSING (IN KIDNEY ISOFORM).
FT VASAPLIC 775 783 MISSING (IN KIDNEY ISOFORM).
FT VASAPLIC 609 1137 MISSING (IN FETAL BRAIN ISOFORM).
FT MDTAGEN 1178 1178 R->A: 2.5-FOLD REDUCTION IN CLEAVAGE.
SQ SEQUENCE 1912 AA; 214759 MW; 3AE8C8D32182E26 CRC64;

Query Match 27.18; Score 45; DB 1; Length 1912;
Best Local Similarity 35.7%; Pred. No. 2.1e+02;
Matches 10; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 2 WRSSGFVGFQGHDPVFLMEPMHV 29
DB 438 WKPEEPNGQLQGYR--VYTMPTQHV 463

RESULT 32
FAS_RAT STANDARD: PRT; 2505 AA.
ID PAS_RAT
AC P12785; 064717; 009187; 009190;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DI 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.39;
EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
GN FASN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89240686; PubMed=2717611;
RA Amy C.M., Witkowski A., Naggert J., Williams B., Randhawa Z.,
RA Smith S.,
RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
RT fatty acid synthase."
RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=93075999; PubMed=339331;
RA Beck K.F., Schreglmann R., Stathopoulos I., Klein H., Hoch J.,
RA Schweizer M.,
RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
RT norvegicus."
RL DNA Seq. 2:359-366(1992).
RN [3]
RP SEQUENCE OF 75-2505 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
RX MEDLINE=89128431; PubMed=2915923;
RA Schweizer M., Takabayashi K., Beck K.F., Schreglmann R.;
RT "Rat mammary gland fatty acid synthase: localization of the
RT constituent domains and two functional polyadenylation/termination
RT signals in the cDNA."
RL Nucleic Acids Res. 17:567-586(1989).
RN [4]
RP SEQUENCE OF 2085-2505 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=88087240; PubMed=2891707;

```

```

RA Naggert J., Witkowski A., Mikkelsen J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
RT domain of the rat fatty acid synthetase."
RL J. Biol. Chem. 263:1146-1150(1988).
RN [5]
RP SEQUENCE OF 1921-2324 FROM N.A.
RC TISSUE=Mammary gland; PubMed=3109907;
RX MEDLINE=87246646; PubMed=3109907;
RA Witkowski A., Naggert J., Mikkelsen J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
RT protein and its flanking domains in the mammalian fatty acid
RT synthetase."
RL Eur. J. Biochem. 165:601-606(1987).
CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
CC ACYL CARRIER PROTEIN.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
CC long-chain fatty acid + (N+1) CoA + N Co(2) + 2N NADP(+) +
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
CC acetyl-[acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
CC malonyl-[acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + Co(2) +
CC [acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxy-palmitoyl-[acyl-carrier protein]
CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
CC carrier protein] + oleate.
CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL: M76767; AAA57219.1; -
DR EMBL: X62888; CAA44678.1; -
DR EMBL: X62889; CAA44680.1; -
DR EMBL: X13415; CAA31780.1; -
DR EMBL: X13527; CAA31882.1; -
DR EMBL: J03514; AAA41144.1; -
DR PIR: A30313; XYRTPA.
DR InterPro: IPR001227; AC transferase.
DR InterPro: IPR002085; Adh_zn_family.
DR InterPro: IPR003880; ketoacyl_synth.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00107; adh_zinc.1.
DR Pfam: PF00109; ketoacyl_synth.1.
DR Pfam: PF00580; pp-binding.1.
DR Pfam: PF00698; Acyl-transf.1.
DR Pfam: PF00975; Thioesterase.1.
DR Pfam: PF02801; ketoacyl_synth_C.1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE: PS50075; ACP_DOMAIN; 1.
KW Fatty acid biosynthesis. Multifunctional enzyme. Phosphopantetheine;
KW Hydroxase. Oxidoreductase. Transferase. Lyase; NADP.
KW Pyridoxal phosphate.
KM HYDROLASE.
FT DOMAIN 429 817 ACYL-KETOACYL SYNTHASE.
FT DOMAIN 1629 1857 ACYL AND MALONYL TRANSFERASES.
FT DOMAIN 1858 2113 ENOYL REDUCTASE.
FT DOMAIN BETA-KETOACYL REDUCTASE.

```

```

FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
FT DOMAIN 2202 2505 THIOESTERASE.
FT ACT_SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 581 581 MALONYLTRANSFERASE (BY SIMILARITY).
FT NP_BIND 1665 1682 NADP (ER).
FT BINDING 1658 1686 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT NP_BIND 1765 1780 NADP (Kt).
FT BINDING 2151 2151 PHOSPHOPANETHEINE (BY SIMILARITY).
FT ACT_SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
FT CONFLICT 871 871 S -> P (IN REF. 3).
FT CONFLICT 1967 1968 MW -> IL (IN REF. 5).
FT CONFLICT 2085 2085 C -> P (IN REF. 4).
FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).
FT CONFLICT 2296 2296 Y -> H (IN REF. 1 AND 5).
SQ SEQUENCE 2505 AA; 272647 MW; 5810BC13D37F3114 CRC64;

Query Match
Best Local Similarity 26.3%; Score 45; DB 1; Length 2505;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SSFYGLFGQHPDYVYLMPEP 25
Db 40 WKAGLGLPKRSGKLMKSLFDASFQVHPKQAHMDP 77

RESULT 33
AATM_BOVIN
ID AATM_BOVIN STANDARD; PRT; 430 AA.
AC P12344;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
GN GOT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RA NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Palmisano A.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 30-41.
RX MEDLINE=79191877; PubMed=446759;
RA Capasso S., Garzillo A.M., Marino G., Mazzarella L., Pucci P.,
RA Sanna G.;
RT "Mitochondrial bovine aspartate aminotransferase. Preliminary
RT sequence and crystallographic data."
RL FEBS Lett. 101:351-354(1979).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: 225466; CA80960.1;
DR PIR: A14290; A14290.
DR HSSP: P00508; TAAT.
DR InterPro: IPR004839; AminoTransf./2.
DR InterPro: IPR000796; AsptTransf_sub.
DR InterPro: IPR004838; NtlTransf_1.
DR Pfam: PF00155; aminotran_12; 1.
DR PRINTS: P00799; TRANSAMINASE.
DR PROSITE: PS00105; AA TRANSFER CLASS 1; 1.
KW Transferrase; Amino transferase; Pyridoxal phosphate; Mitochondrion;
KW Transit peptide.
FT TRANSIT 1 29 MITOCHONDRION.
FT CHAIN 30 430 ASPARTATE AMINOTRANSFERASE.
FT BINDING 279 279 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 430 AA; 47513 MW; 16DDF4753820355A CRC64;

Query Match
Best Local Similarity 44.0%; Score 44.5; DB 1; Length 430;
Matches 11; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 3 RSGSFVGLFGQHPDYVYLMPEAW 27
Db 139 RIGASFGLRFFKFRDVF-LPKPTM 162

RESULT 34
AATM_HUMAN
ID AATM_HUMAN STANDARD; PRT; 430 AA.
AC P00505; Q9BNA3;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
GN GOT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RA [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89087454; PubMed=3207426;
RA Pol S., Bousquet-Lemerclier B., Pave-Preux M., Pawlak A., Nalpas B.,
RA Berthelot P., Hanoune J., Barouki R.;
RT "Nucleotide sequence and tissue distribution of the human
RT mitochondrial aspartate aminotransferase mRNA."
RL Biochem. Biophys. Res. Commun. 157:1305-1315(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 30-430.
RX MEDLINE=86026367; PubMed=4052435;
RA Martini F., Angelaccio S., Barria D., Pascarella S., Maras B.,
RA Doonan S., Bossa F.;
RT "The primary structure of mitochondrial aspartate aminotransferase
RT from human heart."
RL Biochim. Biophys. Acta 832:46-51(1985).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


Query Match 26.5%; Score 44; DB 1; Length 170;
Best Local Similarity 37.5%; Pred. No. 23;
Matches 12; Conservative 2; Mismatches 10; Indels 8; Gaps 1;

OY 4 SSSFFVGLFQGHDPVFYLM 27
DB 68 TSSSLKGFDPKGVDAKYNINHLQGLFTDPBW 99

RESULT 37
RCSA_SALTY
ID RCSA_SALTY STANDARD: PRT: 207 AA.
AC 056083;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Colanic acid capsular biosynthesis activation protein A.
GN RCSA OR STY2190.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TY2;
RX MEDLINE=96198173; PubMed=8626298;
RA Virlogues I., Waxin H., Ecobichon C., Lee J.O., Popoff M.Y.;
RT "Characterization of the rcsA and rcsB genes from Salmonella typhi:
rCSB through tva is involved in regulation of Vi antigen
synthesis";
RT J. Bacteriol. 178:1691-1698(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Patrick J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gea P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhimurium CT18";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE
OF THE GENES FOR CAPSULE SYNTHESIS.
CC SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION
OF THE GENES FOR CAPSULE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE LUXR/OHPA FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87687; CA61019.1; -
DR EMBL: AL627272; CAD05730.1; -
DR InterPro: IPR000792; HTH_LuxR.
DR Pfam: PF00196; Gede; 1.
DR PRINTS: PR00038; HTH_LuxR.
DR PRODOM: PD000307; HTH_LuxR; 1.
DR SMART: SM00421; HTH_LuxR; 1.
DR PROSITE: PS00622; HTH_LuxR_FAMILY; 1.
KW Sensory transduction; Transcription regulation; DNA-binding;
KW Activator; Complete proteome.
FT DNA_BIND 155 174 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23274 MW; 4DC34F9501C66999 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 207;
Best Local Similarity 35.0%; Pred. No. 29;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 4 SSSFFVGLFQGHDPVFYLM 23
DB 63 SDSQIKQIINOHDPVLFIV 82

RESULT 38
RCSA_SALTY
ID RCSA_SALTY STANDARD: PRT: 207 AA.
AC P54698;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Colanic acid capsular biosynthesis activation protein A.
GN RCSA OR SMI1982.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SCSG1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Gwiel N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2";
RT Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN-LT2;
RX MEDLINE=97464436; PubMed=9324257;
RA Chisholm K., Fan F., Schoenhals G.J., Kihara M., Macnab R.M.;
RT "The FljO, FljP, FljQ, and FljR proteins of Salmonella typhimurium:
putative components for flagellar assembly";
RT J. Bacteriol. 179:6092-6099(1997).
CC -1- FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE
OF THE GENES FOR CAPSULE SYNTHESIS.
CC SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION
OF THE GENES FOR CAPSULE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE LUXR/OHPA FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008788; AAU20894.1; -
DR EMBL: LA9021; AAB81322.1; -
DR StGene: SG10580; RCSA.
DR InterPro: IPR000792; HTH_LuxR.
DR Pfam: PF00196; Gede; 1.
DR PRODOM: PD000307; HTH_LuxR; 1.
DR PROSITE: PS00622; HTH_LuxR_FAMILY; 1.
KW Sensory transduction; Transcription regulation; DNA-binding;
KW Activator; Complete proteome.
FT DNA_BIND 155 174 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23302 MW; 4DC34F8A11C66999 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 207;
Best Local Similarity 35.0%; Pred. No. 29;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 4 SSSFFVGLFQGHDPVFYLM 23

Db 63 SDSQIKIINQHPDLEIFV 82

```

RESULT 39
ID Y1FE_ECOLI STANDARD; PRT; 301 AA.
AC P32667;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical transport protein y1fE.
GN Y1FE OR B3943.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;

[1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
CC -1- SIMILARITY: BELONGS TO THE EMA TRANSPORTER FAMILY.
CC -1- SIMILARITY: BELONGS TO THE EMA TRANSPORTER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00006; AAC43049.1; ALT_INT.
DR EMBL; AE000468; AAC76925.1; ALT_INT.
DR EcoGene; EG11902; Y1fE.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 128 148 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
SQ SEQUENCE 301 AA; 32865 MW; 1A822058FDA28696 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 301;
Best Local Similarity 27.6%; Pred. No. 43;
Matches 8; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 WRSQSSFFVQLFGQHPDVFYIMEPAHV 29
Db 163 SWGASAIYAKRLYARHPRVDLSLTSWQM 191

RESULT 40
ID YEO9_SYNT3 STANDARD; PRT; 326 AA.
AC P73594;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical WD-repeat protein slr1409.
GN SLR1409.

```

```

OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;

[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90907; BAA17638.1; -
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 7.
DR SMART; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS00682; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat; Complete proteome.
FT REPEAT 47 77 WD 1.
FT REPEAT 88 118 WD 2.
FT REPEAT 129 159 WD 3.
FT REPEAT 169 199 WD 4.
FT REPEAT 210 240 WD 5.
FT REPEAT 252 282 WD 6.
SQ SEQUENCE 326 AA; 35759 MW; BD83A2403DE1163D CRC64;

Query Match 26.5%; Score 44; DB 1; Length 326;
Best Local Similarity 42.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 WRSQSSFFVQLFGQHPDVF 20
Db 76 WTKGEMTLGQLGQKPMF 94

RESULT 41
ID ALR_DEIRA STANDARD; PRT; 351 AA.
AC Q9VE3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alanine racemase (EC 5.1.1.1).
GN ALR OR DR1086.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;

[1]
RN SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

```

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium *Deinococcus*
 RT *radiodurans* R1.";
 RL Science 286:1571-1577(1999).
 CC -1- FUNCTION: Provides the D-alanine required for cell wall
 CC biosynthesis (By similarity).
 CC -1- CATALYTIC ACTIVITY: L-alanine = D-alanine.
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- PATHWAY: Along with D-alanine-D-alanine ligase, it makes up the
 CC D-alanine branch of the peptidoglycan biosynthetic route.
 CC -1- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE001958; AAF10657.1; ALT-INIT.
 CC HSSP: P10724; 1BD0.
 CC TIGR: DR1086;
 CC InterPro: IPR000821; Ala_racemase.
 CC Pfam: PF00842; Ala_racemase. 1.
 CC PRINTS: PR00992; ALARACEMASE.
 CC TIGRPM: TIGR00492; alt. 1.
 CC PROSITE: PS00395; ALANINE_RACEMASE. 1.
 CC Isomerase; Pyridoxal phosphate; Cell wall; Peptidoglycan synthesis;
 CC Complete proteome.
 CC ACT_SITE 34 CATALYTIC BASE SPECIFIC TO D-ALANINE (BY
 CC SIMILARITY).
 CC ACT_SITE 34 CATALYTIC BASE SPECIFIC TO L-ALANINE (BY
 CC SIMILARITY).
 CC BINDING 34 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC FT SEQUENCE 351 AA; 37192 MW; F924ABD662E2FPA05 CRC64;
 CC SQ
 CC Query Match 26.5%; Score 44; DB 1; Length 351;
 CC Best Local Similarity 50.0%; Pred. No. 50;
 CC Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 CC
 CC QY 5 GGSFVGOLFQGHDPVYFL 22
 CC 1 11:1 111:1
 CC DB 41 GMELVGRLLAAHPDVWGL 58
 CC
 CC RESULT 42
 CC SEQL_XENLA
 CC ID SEQL_XENLA STANDARD; PRT; 481 AA.
 CC AC P58003;
 CC DT 16-OCT-2001 (Rel. 40; Created)
 CC DT 16-OCT-2001 (Rel. 40; Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40; Last annotation update)
 CC DE Sestrin 1 (p53-regulated protein PA26) (XPA26).
 CC GN SEST1 OR PA26.
 CC OS *Xenopus laevis* (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 CC OC Xenopodinae; *Xenopus*.
 CC NCBI_TaxID=8355;
 CC RN
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=21095182; PubMed=1165487;
 CC RA Hikasa H., Taira M.;
 CC RT "A *Xenopus* homolog of a human p53-activated gene, PA26, is
 CC specifically expressed in the notochord.";
 CC RL Mech. Dev. 100:309-312(2001).
 CC CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: AT ZYGOTIC STAGE; SPECIFICALLY EXPRESSED IN
 CC THE NOTOCHORD. MATERNAL TRANSCRIPTS ARE DETECTED AT CLEAVAGE
 CC STAGES AND REDUCED DURING GASTRULATION.

CC -1- SIMILARITY: BELONGS TO THE SESTRIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB048259; BAB33008.1;
 CC DR Nuclear protein.
 CC KW SEQUENCE 481 AA; 55459 MW; 8C175801BBFAFE9D CRC64;
 CC SQ
 CC Query Match 26.5%; Score 44; DB 1; Length 481;
 CC Best Local Similarity 36.7%; Pred. No. 70;
 CC Matches 11; Conservative 6; Mismatches 9; Indels 4; Gaps 2;
 CC
 CC QY 1 SMRS-GGSFVGOLFQGHDPVYFLMEPAWV 29
 CC 1 11:1 111:1
 CC DB 344 SWEHGYSLVNRLY---PDVGLLDEKFTI 370
 CC
 CC RESULT 43
 CC UVRC_HAEIN
 CC ID UVRC_HAEIN STANDARD; PRT; 609 AA.
 CC AC P44489;
 CC DT 01-NOV-1995 (Rel. 32; Created)
 CC DT 01-NOV-1995 (Rel. 32; Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41; Last annotation update)
 CC DE Excinuclease ABC subunit C.
 CC GN UVRC OR HI0057.
 CC OS *Haemophilus influenzae*.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC CC *Haemophilus*.
 CC NCBI_TaxID=727;
 CC RN
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN=Rd / KM20 / ATCC 51907;
 CC MEDLINE=95350630; PubMed=7542800;
 CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 CC Kesteven K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 CC McKenney K., Saiton G., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 CC Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 CC Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 CC Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghagen N.S.M.,
 CC Greth C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 CC Venter J.C.;
 CC RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 CC Rd.";
 CC RL Science 269:496-512(1995).
 CC CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
 CC UVRA-UVRS COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
 CC NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRS AND UVRC.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U32691; AAC21735.1;
 CC DR HSSP: P07025; 1E52.
 CC DR TIGR: HI0057;
 CC DR InterPro: IPR003583; HHH_1.

```

DR InterPro: IPR000445; HhH.
DR InterPro: IPR001943; UVR/C.
DR InterPro: IPR004791; UVR/C.
DR InterPro: IPR001162; UVR/C.
DR InterPro: IPR000305; UVR/N.
DR Pfam: PF00633; HhH. 2.
DR Pfam: PF01541; Excl_endo_N; 1.
DR Pfam: PF02151; UVR; 1.
DR ProDom: PD005870; UVR_C; 1.
DR SMART: SM00465; GYC; 1.
DR SMART: SM00278; HhH; 1.
DR TIGRFAMs: TIGR00194; uvrC; 1.
DR PROSITE: PS50151; UVR; 1.
DR SOS response: Excision nuclease: DNA repair, Complete proteome.
KW DOMAIN 201 236 UVR.
SQ SEQUENCE 609 AA; 69609 MW; 049DD9212349017A CRC64;

Query Match 26.5%; Score 44; DB 1; Length 609;
Best Local Similarity 41.2%; Pred. No. 89;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRGSFVQLFGQHP 17
  |||  |||  |||  |||
Db 114 SYRGSKKFAGEYFGPP 130

RESULT 44
S6AB_MOUSE STANDARD; PRT; 627 AA.
ID S6AB_MOUSE
AC P31650;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sodium- and chloride-dependent GABA transporter 4 (GAT4).
GN SLC6A11 OR GABT4 OR GAT-4 OR GAT4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9311969; PubMed=8420981;
RA Liu Q.-R., Lopez-Coquera B., Mandiyan S., Nelson H., Nelson N.;
RT "Molecular characterization of four pharmacologically distinct gamma-
RT aminobutyric acid transporters in mouse brain.";
RL J. Biol. Chem. 268:2106-2112(1993).
CC -!- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS. HAS A KM
CC OF 0.8 MUM FOR GABA; CAN ALSO TRANSPORT BETA-ALANINE (KM=99 MUM)
CC AND TAURINE (KM=1.4 MUM).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L04662; -!- NOT_ANNOTATED_CDS.
CC PIR: B44409; B44409.
CC MGD: MGI:95630; Gabt4.
CC InterPro: IPR000175; Na/ntran_symport.
CC Pfam: PF00209; SNF; 1.
CC PRINTS: PR00176; NANEUSMPORT.
CC ProDom: PD000448; Na/ntran_symport; 1.
CC PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
CC PROSITE: PS00754; NA_NEUROTRAN_SYM_2; 1.

```

```

DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
KW Symport; Multigene family.
FT DOMAIN 1 53 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 54 74 1 (POTENTIAL).
FT TRANSMEM 82 101 2 (POTENTIAL).
FT TRANSMEM 126 146 3 (POTENTIAL).
FT DOMAIN 147 220 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 221 239 4 (POTENTIAL).
FT TRANSMEM 248 265 5 (POTENTIAL).
FT TRANSMEM 301 318 6 (POTENTIAL).
FT TRANSMEM 330 351 7 (POTENTIAL).
FT TRANSMEM 384 403 8 (POTENTIAL).
FT TRANSMEM 433 451 9 (POTENTIAL).
FT TRANSMEM 468 488 10 (POTENTIAL).
FT TRANSMEM 509 528 11 (POTENTIAL).
FT TRANSMEM 548 566 12 (POTENTIAL).
FT DOMAIN 567 627 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 627 AA; 69888 MW; E6D4E45FC92ACB7 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 627;
Best Local Similarity 41.2%; Pred. No. 92;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 2 WR-----SGGSFVQLFGQHPDVFYIMEPAWY 29
  |||  |||  |||  |||  |||  |||
Db 117 WRVRCPLFEGIGYATQVIEAHLVYIIILAMVI 150

RESULT 45
S6AB_RAT STANDARD; PRT; 627 AA.
ID S6AB_RAT
AC P31647;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Sodium- and chloride-dependent GABA transporter 3.
GN SLC6A11 OR GABT3 OR GAT-3 OR GAT-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=93016029; PubMed=1400419;
RA Borden L.A., Smith K.E., Hartig P.R., Brancheck T.A., Weinschenk R.L.;
RT "Molecular heterogeneity of the gamma-aminobutyric acid (GABA)
RT transport system. Cloning of two novel high affinity GABA
RT transporters from rat brain.";
RL J. Biol. Chem. 267:21098-21104(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92360310; PubMed=1497897;
RA Clark J.A., Deutch A.Y., Gallipoli P.Z., Amara S.G.;
RT "Functional expression and CNS distribution of a beta-alanine-
RT sensitive neuronal GABA transporter.";
RL Neuron 9:337-348(1992).
CC -!- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: BRAIN AND RETINA.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
CC FAMILY (SNP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```



```

RT gastric pathogen Helicobacter pylori." ;
RL Nature 397:176-180(1999) .
CC
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + H(2)O = AMP +
CC phosphoenolpyruvate + phosphate.
CC -1- PATHWAY: ESSENTIAL STEP IN GLUCONEOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -1- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isd-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; AE001450; AAD05690.1. -
DR InterPro; IPR000121; pep_utilizers.
DR InterPro; IPR002192; PPK_N_term.
DR Pfam; PF00391; pep_utilizers; 1.
DR Pfam; PF01326; PPK_N_term; 1.
DR Pfam; PF02896; pep_utilizers; 1.
DR ProDom; PD00940; pep_utilizers; 1.
DR PROSITE; PS00370; pep_ENZYMES_PHOS_SITE; 1.
DR PROSITE; PS00742; pep_ENZYMES_2; 1.
DR Transfaser; Kinase; ATP-binding; 1.
KW phosphorylation, complete proteome,
KW MOD_RES 430 PHOSPHORYLATION (BY SIMILARITY).
FT
SQ SEQUENCE 812 AA; 91291 MW; A088D459B3E47512 CRC64;
QY
Query Match 26.5%; Score 44; DB 1; Length 812;
Best Local Similarity 36.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
DB 220 AMGLGENVGGIT--NPDEFYFKP 242
1 SWRSGSFEVGFQHPDVFYLMPE 25
:|::|::|::|::|
DB 220 AMGLGENVGGIT--NPDEFYFKP 242

RESULT 48
PSSA_HELPY
ID PSSA_HELPY STANDARD: PRT; 812 AA.
AC P56070;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphoenolpyruvate synthase (EC 2.7.9.2) (Pyruvate,water dikinase)
DE (pep synthase).
DE PSSA OR HP0121.
GN
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Adams H.G., Glodek A.,
RA Letteney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,
RA Berg D.E., Gocayne J.D., Utechtack T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Kalp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori." ;
RL Nature 388:539-547(1997) .
CC CATALYTIC ACTIVITY: ATP + pyruvate + H(2)O = AMP +
CC phosphoenolpyruvate + phosphate.
CC -1- PATHWAY: ESSENTIAL STEP IN GLUCONEOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -1- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000534; AAD07191.1; -
CC TIGR; HP0121; -
CC InterPro; IPR000121; PEP_utilizers.
CC InterPro; IPR002192; PDPK_N term.
CC Pfam; PF00391; PEP_utilizers; 1.
CC Pfam; PF01326; PDPK_N term; 1.
CC Pfam; PF02866; PEP_utilizers_C; 1.
CC ProDom; PD000940; PEP_utilizers; 1.
CC PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; 1.
CC TRANSFERASE; KINASE; ATP-binding; Phosphorylation; Complete proteome.
CC MOD_RES 430 430 PHOSPHORYLATION (BY SIMILARITY).
CC SEQUENCE 812 AA; 91273 MW; 56D66703B134AA3 CRC64;
CC -----
CC
CC Query Match 26.5%; Score 44; DB 1; Length 812;
CC Best Local Similarity 36.0%; Pred. No. 1.2e+02;
CC Matches 9; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
CC
CC 1 SWRSSSFVQGLFQGHDPVEYLMKP 25
CC :| | :| | :| | | | :|
CC 220 AMGLGENVVGIT--NPDEFYVFKP 242
CC -----
CC
CC RESULT 49
CC V326_BPMD2 STANDARD; PRT; 836 AA.
CC ID V326_BPMD2 STANDARD; PRT; 836 AA.
CC AC 064220;
CC DT 15-DEC-1998 (Rel. 37, Created)
CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC DE Minor tail protein GP26.
CC GN 26.
CC OS Mycobacteriophage D29.
CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
CC unclassified Siphoviridae.
CC OX NCBI_TaxID=283369;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=96300335; PubMed=9636706;
CC RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
CC RT "Genome structure of mycobacteriophage D29: implications for phage
CC evolution.";
CC RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022214; AAC18467.1; -
CC INTL_MET 0 BY SIMILARITY
CC SEQUENCE 836 AA; 86572 MW; 528D12ED0FC92D6 CRC64;
CC -----
CC
CC Query Match 26.5%; Score 44; DB 1; Length 836;
CC Best Local Similarity 57.9%; Pred. No. 1.2e+02;
CC Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
CC
CC 1 SWRSG-----SSTVQGLFQGL 15
CC :| | | | | | | | | | :|
CC 560 SWSGSGVQGVSDFVQGLPKP 578
CC -----

```

```

RESULT 50
POLR ASGVP
ID POLR ASGVP STANDARD; PRT: 2105 AA.
AC P36309;
DT 01-JUN-1994 (rel. 29, Created)
DT 01-JUN-1994 (rel. 29, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Genome polyprotein [contains: RNA replicase (EC 2.7.7.48); Helicase;
DE Coat protein].
OS Apple stem grooving virus (strain P-209) (ASGV).
OC Viruses; ssRNA positive strand viruses, no DNA stage; Capilliovirus.
OX NCBI_TaxID=36402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93033164; PubMed=1413530;
RA Yoshikawa N., Sasaki E., Kato M., Takahashi T.;
RT "The nucleotide sequence of apple stem grooving capilliovirus genome.";
RL Virology 191:98-105(1992).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- PTM: THE N-TERMINAL OF THE COAT PROTEIN IS BLOCKED.
CC -1- MISCELLANEOUS: THE COAT PROTEIN IS LOCATED IN THE CARBOXY-TERMINAL
CC REGION OF THIS POLYPROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14985; BA03639.1; -.
DR PIR: A44059; A44059.
DR MEROPS: C35.001; -.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR DR PF01443; Viral_helicase1; 1.
KW RNA-directed RNA polymerase; Transferase; Polyprotein; ATP-binding;
KW Coat protein; Helicase.
FT NP_BIND 781 788 ATP (POTENTIAL).
FT DOMAIN 1364 1453 POLYMERASE ACTIVE SITE.
SQ SEQUENCE 2105 AA; 241240 MW; 331291A5FA137131 CRC64;

```

Query Match

Best Local Similarity 33.3%; Pred. No. 3.3e+02;

Matches 9; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

QY 1 SWRSSSFGQLFGQHPDYFLMEPAW 27

DB 1675 SWLGSKQCQCTLEAWKGESELYMEPAW 1701

Search completed: February 20, 2003, 13:32:24
Job time : 41.4286 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:29:11 ; Search time 41.4286 Seconds

(without alignments)
144.233 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166

Sequence: 1 SMRSGSSFWGQLFGHPDYFLMEPAWV 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	166	100.0	09Y5R3	09Y5R3 homo sapien
2	166	100.0	09R1I1	09R1I1 mus musculu
3	166	100.0	09WDE5	09WDE5 mus musculu
4	160	96.4	09GZK3	09GZK3 homo sapien
5	160	96.4	09GUP4	09GUP4 mus musculu
6	157	94.6	09GZS9	09GZS9 homo sapien
7	134	80.7	09EP78	09EP78 mus musculu
8	134	80.7	09NS84	09NS84 mus musculu
9	134	80.7	09S67	09S67 homo sapien
10	131	78.9	09NBO0	09NBO0 mus musculu
11	124	74.7	09UBS5	09UBS5 homo sapien
12	124	74.7	09Y6F2	09Y6F2 homo sapien
13	124	74.7	09B276	09B276 mus musculu
14	124	74.7	09Y4C5	09Y4C5 homo sapien
15	117	70.5	09ECC0	09ECC0 mus musculu
16	116	69.9	043916	043916 homo sapien

17	108	65.1	441	13	093403	093403 torped cal
18	101	60.8	472	11	088199	088199 mus musculu
19	101	60.8	472	11	09QZ12	09QZ12 rattus norv
20	101	60.8	475	4	075099	075099 homo sapien
21	71	42.8	315	5	09VWC3	09VWC3 drosophila
22	71	42.8	486	5	095TN9	095TN9 drosophila
23	57	34.3	46	3	08TG14	08TG14 phaeosphaer
24	57	34.3	1570	16	08YWC4	08YWC4 anabaena sp
25	53	31.9	277	16	099VX8	099VX8 staphylococ
26	53	31.9	277	16	0932E3	0932E3 staphylococ
27	53	31.9	513	8	05S60	05S60 buchiue dac
28	53	31.9	622	2	007653	007653 cellvldrio
29	52.5	31.6	187	17	082VJ3	082VJ3 pyrobaculum
30	52	31.3	46	3	09C409	09C409 phaeosphaer
31	52	31.3	46	3	08TG16	08TG16 phaeosphaer
32	52	31.3	518	16	0913V7	0913V7 pseudomonas
33	52	31.3	614	16	08U1V4	08U1V4 agrobacteri
34	52	31.3	641	16	08YTS6	08YTS6 anabaena sp
35	51	30.7	46	3	09EX13	09EX13 phaeosphaer
36	51	30.7	46	3	09EX15	09EX15 phaeosphaer
37	51	30.7	237	2	09XBL3	09XBL3 bacillus ce
38	51	30.7	263	2	09L3S3	09L3S3 anabaena sp
39	51	30.7	341	10	093WY7	093WY7 securigera
40	51	30.7	681	2	08YR21	08YR21 burkholderi
41	50	30.1	158	11	0925D2	0925D2 rattus norv
42	50	30.1	239	2	08YWG5	08YWG5 pseudomonas
43	50	30.1	678	2	09JN46	09JN46 rhodobacter
44	50	30.1	1051	10	09LXN4	09LXN4 arabidopsis
45	50	30.1	1645	5	099140	099140 drosophila
46	50	30.1	4097	5	09VZ03	09VZ03 drosophila
47	49.5	29.6	634	16	091540	091540 pseudomonas
48	49	29.5	265	4	09NV23	09NV23 homo sapien
49	49	29.5	318	4	09NWT1	09NWT1 homo sapien
50	49	29.5	891	5	095WT2	095WT2 manduca sex
51	49	29.5	4063	5	002425	002425 caenorhabdi
52	48.5	29.2	140	8	09B222	09B222 stenaelius
53	48.5	29.2	361	5	023284	023284 caenorhabdi
54	48.5	29.2	534	16	0910N8	0910N8 streptomyc
55	48.5	29.2	3908	5	09BR91	09BR91 strongyloce
56	48	28.9	245	16	050457	050457 mycobacteri
57	48	28.9	443	10	024532	024532 volvox cart
58	48	28.9	460	16	0916R7	0916R7 pseudomonas
59	48	28.9	518	5	09XU16	09XU16 caenorhabdi
60	48	28.9	581	15	09QBZ3	09QBZ3 simian foam
61	48	28.9	592	16	098H21	098H21 rhizobium l
62	48	28.9	739	13	09PT31	09PT31 oncorhynch
63	48	28.9	739	13	09Y908	09Y908 salmo salar
64	48	28.9	783	13	090Y91	090Y91 oryzias lat
65	48	28.9	817	5	09V838	09V838 drosophila
66	48	28.9	879	13	090W03	090W03 oryzias lat
67	48	28.9	1049	10	09ZNR9	09ZNR9 volvox cart
68	48	28.9	1081	5	09U631	09U631 drosophila
69	48	28.9	1083	5	09VIT0	09VIT0 drosophila
70	48	28.9	3579	5	09VSN8	09VSN8 drosophila
71	47.5	28.6	495	16	083591	083591 treponema p
72	47	28.3	46	3	096X14	096X14 phaeosphaer
73	47	28.3	176	17	09YFK8	09YFK8 aeropyrum p
74	47	28.3	247	16	0968V3	0968V3 rhizobium l
75	47	28.3	262	2	044188	044188 agrobacteri
76	47	28.3	276	5	023526	023526 dictyostell
77	47	28.3	276	5	0921P3	0921P3 rhizobium m
78	47	28.3	284	16	09PEH9	09PEH9 xyella fas
79	47	28.3	288	17	09HPY7	09HPY7 halobacteri
80	47	28.3	296	16	08UB65	08UB65 agrobacteri
81	47	28.3	345	15	0921W7	0921W7 rhizobium m
82	47	28.3	346	16	098045	098045 streptococ
83	47	28.3	349	16	09CJ08	09CJ08 lactococcc
84	47	28.3	363	3	019221	019221 caenorhabdi
85	47	28.3	381	5	001660	001660 caenorhabdi
86	47	28.3	399	13	096SE8	096SE8 gallus galli
87	47	28.3	444	16	08UAP4	08UAP4 agrobacteri
88	47	28.3	456	11	09CVY9	09CVY9 mus musculu
89	47	28.3	465	16	074491	074491 synechocyst

90	47	28.3	469	2	Q9FDT2	Q9Fdt2 shewanella	153	45	27.1	399	11	Q8R0U4	Q8R0u4 mus musculus
91	47	28.3	478	11	Q8VC10	Q8VC10 mus musculus	154	45	27.1	417	10	Q9SFE3	Q9Sfe3 arabidopsis
92	47	28.3	478	11	Q8R430	Q8R430 rattus norv	155	45	27.1	421	10	Q9CAP2	Q9Cap2 arabidopsis
93	47	28.3	496	11	Q9D2D2	Q9D2d2 mus musculu	156	45	27.1	434	16	Q9CUL3	Q9Cul3 pasteurella
94	47	28.3	678	10	Q9C5H9	Q9C5h9 arabidopsis	167	45	27.1	465	16	Q9KRX6	Q9Krx6 vibrio chol
95	47	28.3	701	16	Q9HW59	Q9HW59 pseudomonas	168	45	27.1	485	2	Q8V048	Q8V048 escherichia
96	47	28.3	1008	10	Q9AR83	Q9AR83 linum usita	169	45	27.1	511	8	Q9W0V5	Q9W0V5 olyra latif
97	47	28.3	1075	10	Q9ARC2	Q9ARC2 linum usita	170	45	27.1	622	13	Q91503	Q91503 torpede mar
98	47	28.3	1075	10	Q9AR83	Q9AR83 linum usita	171	45	27.1	648	16	Q69976	Q69976 streptomyce
99	47	28.3	1340	10	Q9LNG5	Q9LNG5 arabidopsis	172	45	27.1	669	3	Q9P392	Q9P392 neurospora
100	47	28.3	3013	16	Q8ZAV3	Q8ZAV3 yersinia pe	173	45	27.1	707	5	Q8S0X2	Q8S0x2 encephalito
101	46.5	28.0	121	2	Q9AMU8	Q9AMU8 bradyrhizob	174	45	27.1	711	5	Q18098	Q18098 caenorhabdi
102	46.5	28.0	163	8	Q9TCU8	Q9TCU8 anisakis si	175	45	27.1	943	4	Q9UQC9	Q9Uqc9 homo sapien
103	46.5	28.0	163	8	Q9TCU8	Q9TCU8 hippostrong	176	45	27.1	943	4	Q9UQC9	Q9Uqc9 homo sapien
104	46.5	28.0	302	16	Q9H0V8	Q9H0v8 homo sapien	177	45	27.1	1051	11	Q9ES34	Q9ES34 mus musculus
105	46.5	28.0	306	16	Q9AMN7	Q9AMn7 rhizobium 1	178	45	27.1	1254	11	Q8VEV0	Q8VEV0 mus musculus
106	46.5	28.0	313	16	Q9KX98	Q9KX98 streptomyce	179	45	27.1	1279	16	Q9C187	Q9C187 lactococcus
107	46.5	28.0	421	2	Q9AM36	Q9AM36 desulfotolvir	180	45	27.1	1654	16	Q8XY86	Q8XY86 anabaena sp
108	46.5	28.0	543	2	Q9F7E4	Q9F7e4 acinetobact	181	45	27.1	1894	11	Q64487	Q64487 mus musculus
109	46.5	28.0	543	2	Q9R2F5	Q9R2f5 acinetobact	182	45	27.1	2109	12	Q83743	Q83743 mus musculus
110	46.5	28.0	534	16	Q9PC78	Q9PC78 xyella fas	183	45	27.1	2109	12	Q63667	Q63667 beet necrot
111	46.5	28.0	708	5	Q8T6B8	Q8T6b8 dictyostell	184	45	27.1	2117	12	Q91R44	Q91R44 beet soil-b
112	46.5	28.0	717	4	Q9MX10	Q9MX10 homo sapien	185	45	27.1	2504	11	Q9EQR0	Q9EQR0 mus musculus
113	46.5	28.0	820	10	Q9SXB4	Q9SXB4 arabidopsis	186	45	27.1	2505	11	Q63577	Q63577 rattus norv
114	46.5	28.0	941	4	Q9P2S6	Q9P2s6 homo sapien	187	45	27.1	2944	11	Q63870	Q63870 mus musculus
115	46.5	28.0	46	3	Q8TGM3	Q8TGM3 phaeosphaer	188	44.5	26.8	401	16	Q9CB95	Q9CB95 mycobacteri
116	46.5	28.0	46	3	Q8TGM3	Q8TGM3 apis mellif	189	44.5	26.8	430	16	Q8U556	Q8U556 agrobacteri
117	46.5	28.0	127	5	Q9N1M1	Q9N1m1 apis mellif	190	44.5	26.8	469	10	Q80456	Q80456 arabidopsis
118	46.5	28.0	175	10	Q9SV53	Q9SV53 arabidopsis	191	44.5	26.8	483	5	Q9VFS7	Q9Vfs7 delnoccocus
119	46.5	28.0	225	8	Q9XMB8	Q9XMB8 ceratilis c	192	44.5	26.8	540	16	Q9R0B8	Q9R0B8 thermomae
120	46.5	28.0	228	10	Q9FRU3	Q9FRu3 glycine max	193	44.5	26.8	567	5	Q8T9L5	Q8T9L5 drosophila
121	46.5	28.0	235	5	Q97421	Q97421 drosophila	194	44.5	26.8	583	16	Q8RCR3	Q8RCr3 thermomae
122	46.5	28.0	235	5	Q9W4X1	Q9W4x1 drosophila	195	44.5	26.8	583	16	Q8RCR3	Q8RCr3 thermomae
123	46.5	28.0	276	2	Q52431	Q52431 pseudomonas	196	44.5	26.8	652	2	Q93H25	Q93H25 streptomyce
124	46.5	28.0	277	9	Q8SCS5	Q8SCS5 pseudomonas	197	44.5	26.8	674	10	Q9F1B4	Q9F1B4 caenorhabdi
125	46.5	28.0	291	16	Q914U7	Q914u7 pseudomonas	198	44.5	26.8	687	5	Q23387	Q23387 agrobacteri
126	46.5	28.0	364	16	Q9RZB3	Q9RZb3 delnoccocus	199	44.5	26.8	1266	16	Q8UGR8	Q8UGr8 agrobacteri
127	46.5	28.0	376	8	Q9MSD6	Q9MSd6 limonium mo	200	44.5	26.8	1546	12	Q8ORC8	Q8ORc8 white spot
128	46.5	28.0	378	8	Q9MSD9	Q9MSd9 limonium ru	201	44.5	26.8	1564	12	Q91LA2	Q91La2 white spot
129	46.5	28.0	418	10	Q9A6A4	Q9A6a4 arabidopsis	202	44.5	26.8	1867	3	Q9P8M4	Q9P8m4 blumeria gr
130	46.5	28.0	426	27.7	Q9LZB3	Q9LZb3 arabidopsis	203	44.5	26.8	2207	5	Q9U0Y2	Q9U0y2 leishmania
131	46.5	28.0	510	8	Q9MW06	Q9MW06 lithacine p	204	44	26.5	150	11	Q70337	Q70337 mus musculu
132	46.5	28.0	513	8	Q9T1A4	Q9T1a4 dinebra ret	205	44	26.5	158	11	Q99L97	Q99L97 mus musculu
133	46.5	28.0	537	16	Q9CENZ	Q9CENZ lactococcus	206	44	26.5	158	16	Q92K71	Q92K71 rhizobium m
134	46.5	28.0	591	5	Q9M9J1	Q9M9j1 leishmania	207	44	26.5	218	10	Q9ZV89	Q9ZV89 arabidopsis
135	46.5	28.0	764	16	Q9KCI0	Q9KCI0 bacillomys ha	208	44	26.5	230	17	Q8TZ07	Q8TZ07 pyrococcus
136	46.5	28.0	6048	2	Q93H87	Q93H87 streptomyce	209	44	26.5	238	10	Q8RU78	Q8RU78 oiyza sativ
137	46.5	28.0	136	10	Q8S5H8	Q8S5h8 cryza sativ	210	44	26.5	254	16	P74150	P74150 synchocyst
138	46.5	28.0	299	5	Q96603	Q96603 plodia inte	211	44	26.5	265	11	Q8R197	Q8R197 mus musculu
139	46.5	28.0	316	16	Q9KXU1	Q9KXU1 cytosfridium	212	44	26.5	267	4	Q8W4H1	Q8W4H1 homo sapien
140	46.5	28.0	341	5	Q17972	Q17972 caenorhabdi	213	44	26.5	283	11	Q9CRR6	Q9CRR6 mus musculu
141	46.5	28.0	493	5	Q8TA03	Q8TA03 dictyostell	214	44	26.5	296	11	Q8V8N9	Q8V8N9 methanosarc
142	46.5	28.0	622	16	Q82KM1	Q82KM1 salmoneilla	215	44	26.5	305	17	Q8TRD6	Q8TRd6 methanosarc
143	46.5	28.0	708	3	Q9C107	Q9C107 aspergillus	216	44	26.5	312	2	Q6A338	Q6A338 vibrio chol
144	46.5	28.0	809	16	Q8U4L9	Q8U4l9 agrobacteri	217	44	26.5	312	2	Q9AGX3	Q9AGx3 vibrio chol
145	46.5	28.0	844	16	P73363	P73363 synchocyst	218	44	26.5	312	16	Q9KXR7	Q9KXR7 vibrio chol
146	46.5	28.0	4547	5	Q9W343	Q9W343 drosophila	219	44	26.5	312	16	Q8X764	Q8X764 escherichia
147	46.5	28.0	103	16	Q8XUD3	Q8XUD3 ralslonia s	220	44	26.5	334	6	Q97724	Q97724 felis silve
148	46.5	28.0	127	5	Q9N1M0	Q9N1m0 trichogramma	221	44	26.5	362	16	Q8ZCJ8	Q8ZCj8 yersinia pe
149	46.5	28.0	138	5	Q8W0I3	Q8W0i3 osteria	222	44	26.5	364	17	Q9UY73	Q9UY73 pyrococcus
150	46.5	28.0	214	16	P73454	P73454 caulobacter	223	44	26.5	383	16	Q8Y373	Q8Y373 ralslonia s
151	46.5	28.0	218	16	Q9ASK1	Q9ASk1 caulobacter	224	44	26.5	394	16	Q9TWS8	Q9TWS8 clostridium
152	46.5	28.0	225	16	Q8H236	Q8H236 pseudomonas	225	44	26.5	405	17	Q28871	Q28871 archaeoglob
153	46.5	28.0	239	16	Q9ALU6	Q9ALU6 streptococc	226	44	26.5	413	9	Q37930	Q37930 bacterioph
154	46.5	28.0	246	16	Q8YVM6	Q8YVM6 anabaena sp	227	44	26.5	414	16	Q977L2	Q977L2 lotus japon
155	46.5	28.0	247	2	Q44300	Q44300 aeromonas s	228	44	26.5	418	10	Q83781	Q83781 lotus japon
156	46.5	28.0	259	2	Q8RR78	Q8RR78 acetobacter	229	44	26.5	439	16	Q8U6J7	Q8U6j7 agrobacteri
157	46.5	28.0	313	2	Q9EXV3	Q9EXv3 rhizobium m	230	44	26.5	445	10	Q22471	Q22471 oiyza sativ
158	46.5	28.0	336	16	Q8XOL6	Q8XOL6 ralslonia s	231	44	26.5	445	16	Q8XXL4	Q8XXL4 ralslonia s
159	46.5	28.0	338	16	Q9WZ03	Q9WZ03 thermotoga	232	44	26.5	449	17	Q9HM03	Q9HM03 thermoplasm
160	46.5	28.0	370	13	Q918K8	Q918K8 brachydanio	233	44	26.5	449	17	Q97CS7	Q97CS7 thermoplasm
161	46.5	28.0	370	16	Q9HY00	Q9HY00 pseudomonas	234	44	26.5	457	16	Q99SB8	Q99SB8 staphylococ
162	46.5	28.0	399	11	Q9QX69	Q9QX69 rattus sp.	235	44	26.5	460	16	Q67510	Q67510 aquifex aeo

236	44	26.5	471	3	Q9C295	Q9C295	neurospora	309	43	25.9	387	4	Q96FAS	Q96FAS	homo sapien
237	44	26.5	476	17	Q9THC2	Q8thc2	methanosarc	310	43	25.9	387	17	Q57972	Q57972	pyrococcus
238	44	26.5	478	2	Q9KXF0	Q9kxf0	escherichia	311	43	25.9	410	16	Q98250	Q98250	rhizobium 1
239	44	26.5	478	2	Q8VNP7	Q8vnp7	escherichia	312	43	25.9	413	10	Q01849	Q01849	phytophthor
240	44	26.5	478	9	Q9MCT4	Q9mct4	bacterioph	313	43	25.9	427	16	Q8RYV9	Q8ryv9	anabaena sp
241	44	26.5	478	9	Q9A217	Q9a217	bacterioph	314	43	25.9	444	10	Q55744	Q55744	cicer ariet
242	44	26.5	483	9	Q9TIM4	Q9tim4	bacterioph	315	43	25.9	445	10	Q22470	Q22470	oryza sativ
243	44	26.5	483	16	Q8XZ8R	Q8xz8r	escherichia	316	43	25.9	462	4	Q96E28	Q96E28	homo sapien
244	44	26.5	505	5	Q9VFN2	Q9vfn2	drosophila	317	43	25.9	468	8	Q47136	Q47136	epigaea rep
245	44	26.5	509	5	Q15712	Q15712	paramecium	318	43	25.9	469	8	Q9FAR9	Q9far9	sheanella
246	44	26.5	513	8	Q9TIB8	Q9tib8	bouteloua c	319	43	25.9	473	6	Q28164	Q28164	bos taurus
247	44	26.5	513	8	Q95F52	Q95f52	gymnopocon	320	43	25.9	474	16	Q96YIG2	Q96yig2	bruceia me
248	44	26.5	540	2	Q936X7	Q936x7	pseudomonas	321	43	25.9	477	4	Q96S88	Q96s88	homo sapien
249	44	26.5	563	15	Q9KXJ5	Q9kxj5	streptomyces	322	43	25.9	494	4	Q9N2T7	Q9n2t7	homo sapien
250	44	26.5	577	15	Q9QBN9	Q9qbn9	simlan foam	323	43	25.9	495	2	Q52406	Q52406	edwardsiell
251	44	26.5	633	2	Q9AGW3	Q9agw3	pseudomonas	324	43	25.9	502	16	Q8VJN6	Q8vjn6	mycobacteri
252	44	26.5	633	16	Q9ZY93	Q9zy93	rhizobium m	325	43	25.9	512	8	Q9MUY7	Q9muy7	leymus angu
253	44	26.5	638	16	Q9KRB0	Q9kbr0	bacillus ha	326	43	25.9	513	8	Q95F41	Q95f41	tiraphis s
254	44	26.5	723	2	Q9ZNM8	Q9znm8	sphingomona	327	43	25.9	530	13	Q9DFV2	Q9dfv2	colurnix co
255	44	26.5	749	2	Q9EFN9	Q9efn9	uncultured	328	43	25.9	534	4	Q75497	Q75497	homo sapien
256	44	26.5	773	2	Q93PB8	Q93pb8	azorarcus ey	329	43	25.9	572	16	Q9S1S8	Q9s1s8	streptomyces
257	44	26.5	778	13	Q9STP4	Q9stp4	oreochromis	330	43	25.9	672	12	Q8QLR6	Q8qlr6	mamestra co
258	44	26.5	940	10	Q9AV70	Q9av70	oryza sativ	331	43	25.9	678	2	Q54679	Q54679	streptomyces
259	44	26.5	2376	2	Q9F9Z3	Q9f9z3	serratia en	332	43	25.9	694	16	Q8X710	Q8x710	escherichia
260	44	26.5	2509	4	Q16702	Q16702	homo sapien	333	43	25.9	700	16	Q8XIM4	Q8xim4	raistonia s
261	44	26.5	3117	4	Q8WY20	Q8wy20	homo sapien	334	43	25.9	759	11	Q9QMS6	Q9qms6	mus musc
262	44	26.5	6885	4	Q8WYH0	Q8wyh0	homo sapien	335	43	25.9	771	11	Q9QMS5	Q9qms5	mus musc
263	44	26.5	51	2	Q05977	Q05977	synecocyst	336	43	25.9	874	11	Q9QJF8	Q9qjf8	mus musc
264	43.5	26.2	105	16	Q9N8G6	Q9n8g6	rhizobium 1	337	43	25.9	889	10	Q9FK48	Q9fk48	arabidopsis
265	43.5	26.2	148	11	Q9D213	Q9d213	mus musc	338	43	25.9	902	2	Q9JMX0	Q9jmx0	bradyrhizob
266	43.5	26.2	157	16	Q67847	Q67847	aguiex aeo	339	43	25.9	962	4	Q94977	Q94977	anabaena sp
267	43.5	26.2	222	16	Q8XPM4	Q8xpm4	raistonia s	340	43	25.9	1069	16	Q8RYV9	Q8ryv9	pyrococcus
268	43.5	26.2	320	16	Q916V5	Q916v5	pseudomonas	341	43	25.9	1168	16	Q9QNM0	Q9qnm0	staphylococ
269	43.5	26.2	328	16	Q8XAK2	Q8xak2	escherichia	342	43	25.9	1206	11	Q91ZJ6	Q91zj6	mus musc
270	43.5	26.2	357	8	Q9F9X5	Q9f9x5	cancer ante	343	43	25.9	1234	11	Q8VDJ4	Q8vdj4	mus musc
271	43.5	26.2	406	12	Q9E205	Q9e205	cercopithec	344	43	25.9	1345	11	Q9ZU84	Q9zu84	staphyloche
272	43.5	26.2	505	9	Q38341	Q38341	lactococcus	345	43	25.9	1500	10	Q9ZU84	Q9zu84	arabidopsis
273	43.5	26.2	510	8	Q8WHL3	Q8whl3	wolffella	346	43	25.9	1716	11	Q9QMS7	Q9qms7	mus musc
274	43.5	26.2	527	17	Q9H1Q1	Q9h1q1	thermoplasma	347	43	25.9	1908	5	Q9Y8K9	Q9y8k9	drosophila
275	43.5	26.2	564	16	Q8YPL1	Q8ypl1	anabaena sp	348	43	25.9	2970	12	Q56073	Q56073	hepatitis g
276	43.5	26.2	587	2	Q69353	Q69353	rhodococcus	349	42.5	25.6	130	2	Q91UV2	Q91uv2	escherichia
277	43.5	26.2	631	2	Q8YPE5	Q8ype5	pseudomonas	350	42.5	25.6	171	16	Q8U7Q3	Q8u7q3	agrobacteri
278	43.5	26.2	709	10	Q9K4X4	Q9k4x4	zea mays (m	351	42.5	25.6	179	16	Q8YWK3	Q8ywk3	anabaena sp
279	43.5	26.2	797	11	Q9GWH7	Q9gwh7	rattus norv	352	42.5	25.6	210	16	Q9A1P2	Q9a1p2	streptococ
280	43	25.9	83	2	Q47583	Q47583	gallus gall	353	42.5	25.6	240	16	P73822	P73822	synecocyst
281	43	25.9	90	6	Q9N1A0	Q9n1a0	canis fami	354	42.5	25.6	260	10	Q946T6	Q946t6	oryza sativ
282	43	25.9	90	6	Q9N1A0	Q9n1a0	canis fami	355	42.5	25.6	262	16	P74256	P74256	oryza sativ
283	43	25.9	112	2	Q9FDP4	Q9fdp4	staphylococ	356	42.5	25.6	287	17	Q8ZWS8	Q8zws8	pyrococcus
284	43	25.9	126	5	Q9NGO4	Q9ngo4	diarotica	357	42.5	25.6	303	16	Q9RIT6	Q9rit6	streptococ
285	43	25.9	130	13	Q91870	Q91870	gallus gall	358	42.5	25.6	305	16	Q67486	Q67486	aguiex aeo
286	43	25.9	130	13	Q91870	Q91870	gallus gall	359	42.5	25.6	308	16	Q9PKP6	Q9pkp6	chlamydia m
287	43	25.9	147	16	Q8XYQ4	Q8xyq4	anabaena sp	360	42.5	25.6	344	3	Q9YUV7	Q9yuv7	candida alb
288	43	25.9	151	5	Q9Y036	Q9y036	helicoverpa	361	42.5	25.6	389	17	Q8Z2H7	Q8z2h7	pyrobaculum
289	43	25.9	158	10	Q91TQ2	Q91tq2	arabidopsis	362	42.5	25.6	395	12	Q8E208	Q8e208	cercopithec
290	43	25.9	162	16	Q9ZBM0	Q9zbm0	listeria in	363	42.5	25.6	421	5	Q8T917	Q8t917	drosophila
291	43	25.9	162	16	Q9ZBM0	Q9zbm0	listeria in	364	42.5	25.6	453	4	Q9BXJ7	Q9bxj7	homo sapien
292	43	25.9	165	16	Q8YK6M	Q8yk6m	mycobacteri	365	42.5	25.6	509	10	Q913H5	Q913h5	spirodela i
293	43	25.9	170	1	P71810	P71810	methanosarc	366	42.5	25.6	527	16	Q913H5	Q913h5	pseudomonas
294	43	25.9	170	1	P72020	P72020	methanosarc	367	42.5	25.6	529	10	Q940R3	Q940r3	arabidopsis
295	43	25.9	245	16	Q939C6	Q939c6	prochloroco	368	42.5	25.6	574	16	P73758	P73758	synecocyst
296	43	25.9	264	16	Q8XLU8	Q8xlu8	lactococcus	369	42.5	25.6	590	16	Q54525	Q54525	streptococ
297	43	25.9	305	16	Q8R7Z5	Q8r7z5	thermoanaer	370	42.5	25.6	590	16	Q9A1I6	Q9a1i6	streptococ
298	43	25.9	307	2	Q9S270	Q9s270	homo sapien	371	42.5	25.6	622	5	Q8XAW9	Q8xaw9	escherichia
299	43	25.9	308	2	Q93E11	Q93e11	rhizobium 1	372	42.5	25.6	657	16	Q9VBR6	Q9vbr6	drosophila
300	43	25.9	323	8	Q94WB9	Q94wb9	xenisthmus	373	42.5	25.6	686	10	Q8VYG0	Q8vyg0	arabidopsis
301	43	25.9	335	10	Q43760	Q43760	glycine max	374	42.5	25.6	692	10	Q91VY3	Q91vy3	arabidopsis
302	43	25.9	360	16	Q53536	Q53536	mycobacteri	375	42.5	25.6	692	10	Q93ZM6	Q93zm6	arabidopsis
303	43	25.9	370	12	Q82512	Q82512	influenzavi	376	42.5	25.6	823	16	Q93UT8	Q93ut8	neisseria m
304	43	25.9	373	16	Q9RT51	Q9rt51	deinococcus	377	42.5	25.6	892	10	Q82484	Q82484	neisseria m
305	43	25.9	375	16	Q9KUT3	Q9kut3	vibri	378	42.5	25.6	1069	16	Q9JF35	Q9jf35	arabidopsis
306	43	25.9	376	8	Q9TAH8	Q9tah8	cateretia r	379	42.5	25.6	1069	16	Q9JY62	Q9jy62	neisseria m
307	43	25.9	382	3	Q9F8U8	Q9f8u8	candida alb	380	42.5	25.6	1069	16	Q9JY62	Q9jy62	neisseria m
308	43	25.9	383	17	Q8ZV07	Q8zv07	pyrobaculum	381	42.5	25.6	2304	5	Q9W548	Q9w548	drosophila

382	42.5	25.6	2422	5	046083	046083 drosophila	455	42	25.3	746	12	08V7H8	08V7H8 tt virus. o
383	42	25.3	156	11	099JRO	099JRO mus musculus	456	42	25.3	791	12	09LJF5	09LJF5 streptomyc
384	42	25.3	196	16	08U5X0	08U5X0 agrobacteri	457	42	25.3	791	16	09K496	09K496 streptomyc
385	42	25.3	205	11	09CSB0	09CSB0 mus musculus	458	42	25.3	794	17	09TNR0	09TNR0 methanosarc
386	42	25.3	218	6	09SKW3	09SKW3 bos taurus	459	42	25.3	805	2	09LJ83	09LJ83 streptomyc
387	42	25.3	239	4	09BX36	09BX36 homo sapien	460	42	25.3	824	3	09HFU0	09HFU0 neosporea
388	42	25.3	253	1	093679	093679 sulfolobus	461	42	25.3	848	5	025198	025198 hydra atten
389	42	25.3	258	5	09V6J4	09V6J4 drosophila	462	42	25.3	858	15	099137	099137 avian endog
390	42	25.3	264	16	09XA68	09XA68 bacillus ha	463	42	25.3	861	16	006944	006944 synecocyst
391	42	25.3	264	16	08XM43	08XM43 clostridium	464	42	25.3	876	2	09RMV6	09RMV6 staphylococ
392	42	25.3	268	4	096107	096107 homo sapien	465	42	25.3	876	16	099TH2	099TH2 staphylococ
393	42	25.3	279	16	0921M2	0921M2 rhizobium m	466	42	25.3	879	15	P89909	P89909 chimpanzee
394	42	25.3	280	16	09X8M2	09X8M2 streptomyc	467	42	25.3	902	11	08R0Y6	08R0Y6 mus musculu
395	42	25.3	290	2	09ROV3	09ROV3 neisseria m	468	42	25.3	905	5	09XWS2	09XWS2 caenorhabd
396	42	25.3	290	4	09N0U4	09N0U4 homo sapien	469	42	25.3	921	12	08OS33	08OS33 chimpanzee
397	42	25.3	290	16	09X0M6	09X0M6 neisseria m	470	42	25.3	1019	10	P94078	P94078 arbidops
398	42	25.3	290	16	09J122	09J122 neisseria m	471	42	25.3	1094	2	093H78	093H78 streptomyc
399	42	25.3	303	16	093JB6	093JB6 streptomyc	472	42	25.3	1253	6	097566	097566 canis famill
400	42	25.3	327	9	08SDC3	08SDC3 pseudomonas	473	42	25.3	1354	5	076251	076251 drosophila
401	42	25.3	332	4	09UKF7	09UKF7 homo sapien	474	42	25.3	1364	5	09V6K1	09V6K1 drosophila
402	42	25.3	335	4	09H0U3	09H0U3 homo sapien	475	42	25.3	1816	16	09KBB4	09KBB4 bacillus ha
403	42	25.3	339	16	09HUI3	09HUI3 pseudomonas	476	42	25.3	1841	2	09FB25	09FB25 streptomyc
404	42	25.3	345	11	09D790	09D790 mus musculus	477	42	25.3	1896	13	091AJ1	091AJ1 xenopus lae
405	42	25.3	357	17	09HR66	09HR66 halobacteri	478	42	25.3	4513	10	09MBF8	09MBF8 chlamydomon
406	42	25.3	362	17	08U197	08U197 pyrococcus	479	42	25.3	175	10	09C871	09C871 arbidops
407	42	25.3	365	2	09KH45	09KH45 pantoea agg	480	42	25.0	184	10	09LQ28	09LQ28 arbidops
408	42	25.3	382	11	09CSN5	09CSN5 mus musculus	481	42	25.0	237	5	08MRE9	08MRE9 strongyloce
409	42	25.3	387	10	09ZU26	09ZU26 arbidops	482	42	25.0	250	3	094584	094584 schizosacch
410	42	25.3	392	16	P73562	P73562 synecocyst	483	42	25.0	267	2	09FOX0	09FOX0 pseudomonas
411	42	25.3	392	16	09XAL3	09XAL3 caulobacter	484	42	25.0	271	16	08XM06	08XM06 clostridium
412	42	25.3	398	16	0558T0	0558T0 synecocyst	485	42	25.0	282	2	093ME7	093ME7 amycolatops
413	42	25.3	403	5	09RYR5	09RYR5 delnoccocus	486	42	25.0	304	16	08X2N5	08X2N5 escherichia
414	42	25.3	404	5	016752	016752 caenorhabdi	487	42	25.0	305	16	08X4G4	08X4G4 escherichia
415	42	25.3	418	11	09CT22	09CT22 mus musculu	488	42	25.0	327	11	09EOW3	09EOW3 mus musculu
416	42	25.3	423	16	08YPM5	08YPM5 anabaena sp	489	42	25.0	332	11	09EOW2	09EOW2 mus musculu
417	42	25.3	428	16	099SV2	099SV2 staphylococ	490	42	25.0	334	17	09V046	09V046 pyrococcus
418	42	25.3	434	17	08ZMP3	08ZMP3 pyrobaculum	491	42	25.0	346	16	09KVC6	09KVC6 vibrio chol
419	42	25.3	444	10	022402	022402 nicotiana t	492	42	25.0	402	4	09BR75	09BR75 homo sapien
420	42	25.3	448	16	09PLJ3	09PLJ3 chlamydia m	493	42	25.0	425	15	098877	098877 rhizobium l
421	42	25.3	450	2	055109	055109 synecococ	494	42	25.0	425	16	098877	098877 rhizobium l
422	42	25.3	451	2	050652	050652 paracoccus	495	42	25.0	422	2	P72444	P72444 streptococ
423	42	25.3	453	2	P95767	P95767 streptococ	496	42	25.0	422	5	025242	025242 lucilia cup
424	42	25.3	453	2	P72446	P72446 streptococ	497	42	25.0	463	10	09FHV6	09FHV6 arbidops
425	42	25.3	454	16	092B27	092B27 listeria in	498	42	25.0	480	5	025241	025241 lucilia cup
426	42	25.3	454	11	09D311	09D311 mus musculu	499	42	25.0	482	16	09KLS0	09KLS0 vibrio chol
427	42	25.3	478	2	08YDA5	08YDA5 vitreoscill	500	42	25.0	504	16	09K567	09K567 vibrio chol
428	42	25.3	481	2	093Q31	093Q31 salmonella	501	42	25.0	528	10	09SHE3	09SHE3 arbidops
429	42	25.3	492	4	08WUA2	08WUA2 homo sapien	502	42	25.0	533	5	09W1S3	09W1S3 drosophila
430	42	25.3	492	11	09CXG3	09CXG3 mus musculu	503	42	25.0	537	16	09E2K6	09E2K6 streptomyc
431	42	25.3	499	12	036448	036448 fowlpox vir	504	42	25.0	537	16	09NKR6	09NKR6 leishmania
432	42	25.3	501	11	0920M2	0920M2 cavia porce	505	42	25.0	557	2	09KOS3	09KOS3 vibrio chol
433	42	25.3	506	8	003054	003054 triticum ae	506	42	25.0	619	2	045531	045531 bacillus sp
434	42	25.3	512	8	09MUY6	09MUY6 triticum ae	507	42	25.0	630	16	09K9N1	09K9N1 bacillus ha
435	42	25.3	513	2	092401	092401 cornebacte	508	42	25.0	635	5	08U273	08U273 caenorhabd
436	42	25.3	513	8	09MUY2	09MUY2 loudetiopsi	509	42	25.0	642	10	09LEB1	09LEB1 arbidops
437	42	25.3	517	4	09POV9	09POV9 homo sapien	510	42	25.0	651	10	094AY1	094AY1 arbidops
438	42	25.3	520	16	08YCA4	08YCA4 bruceella me	511	42	25.0	724	15	09N6I4	09N6I4 leishmania
439	42	25.3	531	16	08RKH6	08RKH6 viscum albu	512	42	25.0	724	16	09CJU9	09CJU9 pasteurella
440	42	25.3	531	16	08ZG28	08ZG28 yersinia pe	513	42	25.0	730	4	09BRY5	09BRY5 homo sapien
441	42	25.3	557	5	023933	023933 drosophila	514	42	25.0	764	10	P93756	P93756 arbidops
442	42	25.3	561	2	052072	052072 rhodococcus	515	42	25.0	802	16	09A511	09A511 caulobacter
443	42	25.3	571	10	0489B1	0489B1 oryza sativ	516	42	25.0	823	16	09FWM0	09FWM0 oryza sativ
444	42	25.3	577	15	09QBR8	09QBR8 simian foam	517	42	25.0	982	10	09FWM0	09FWM0 oryza sativ
445	42	25.3	577	15	09QBR7	09QBR7 simian foam	518	42	25.0	993	16	09PDJ4	09PDJ4 xyella fas
446	42	25.3	577	15	09QBR6	09QBR6 simian foam	519	42	25.0	1048	3	093K40	093K40 klebsiella
447	42	25.3	577	15	09QBR5	09QBR5 simian foam	520	42	25.0	1245	5	077202	077202 acanthamoeb
448	42	25.3	577	15	09QBR4	09QBR4 simian foam	521	42	25.0	1244	12	086925	086925 aurea virus.
449	42	25.3	577	15	09QBR3	09QBR3 simian foam	522	42	25.0	1363	5	020766	020766 caenorhabd
450	42	25.3	577	15	09QBR2	09QBR2 simian foam	523	42	25.0	1753	5	09BHW9	09BHW9 leishmania
451	42	25.3	583	17	026600	026600 methanobact	524	42	25.0	2336	5	077291	077291 drosophila
452	42	25.3	621	2	093HU3	093HU3 rhodothermu	525	42	25.0	2535	16	08ZDR6	08ZDR6 yersinia pe
453	42	25.3	633	16	08XYX5	08XYX5 ralstonia s	526	42	25.0	2759	5	045614	045614 caenorhabd
454	42	25.3	667	16	08U511	08U511 agrobacteri	527	42	25.0	3102	5	09T2R4	09T2R4 caenorhabd

528	41.5	25.0	5201	5	Q9U479	Q9U479 drosophila	601	41	24.7	376	5	Q9U5J1	Q9U5J1 trypanosoma
529	41.5	25.0	5385	5	Q9VEV3	Q9VEV3 drosophila	602	41	24.7	379	5	Q76124	Q76124 papilio xut
530	41.5	25.0	5496	5	Q9VEV2	Q9VEV2 drosophila	603	41	24.7	380	5	Q9UAW7	Q9UAW7 papilio xut
531	41.5	25.0	8805	5	Q9VEV4	Q9VEV4 drosophila	604	41	24.7	389	16	Q9A412	Q9A412 caulobacter
532	41	24.7	109	5	Q9V864	Q9V864 drosophila	605	41	24.7	390	17	Q8TMD1	Q8TMD1 metanosarc
533	41	24.7	119	5	Q9U4W3	Q9U4W3 plasmodium	606	41	24.7	392	2	Q93CN3	Q93CN3 bacteroides
534	41	24.7	123	10	Q91GCP5	Q91GCP5 oryza sativ	607	41	24.7	392	5	Q8S5T6	Q8S5T6 dictyosteli
535	41	24.7	135	2	Q86381	Q86381 nostoc comm	608	41	24.7	393	16	Q33319	Q33319 mycobacteri
536	41	24.7	136	2	Q86421	Q86421 nostoc sp.	609	41	24.7	400	16	Q8ZEO1	Q8ZEO1 versinia pe
537	41	24.7	138	5	Q81597	Q81597 wolinnella s	610	41	24.7	401	3	Q9UT27	Q9UT27 schizosacch
538	41	24.7	138	5	Q81597	Q81597 plasmodium	611	41	24.7	401	16	Q8XZ8	Q8XZ8 ralsconia s
539	41	24.7	139	5	Q81526	Q81526 neisseria m	612	41	24.7	404	16	Q8XSW0	Q8XSW0 ralsconia s
540	41	24.7	144	2	Q52409	Q52409 neisseria m	613	41	24.7	405	5	Q964F2	Q964F2 crithidia f
541	41	24.7	144	2	Q54609	Q54609 neisseria m	614	41	24.7	411	5	Q18317	Q18317 caenorhadi
542	41	24.7	144	2	Q54638	Q54638 neisseria m	615	41	24.7	416	10	Q80731	Q80731 caenorhadi
543	41	24.7	144	2	Q54558	Q54558 neisseria m	616	41	24.7	435	16	Q97N85	Q97N85 streptococ
544	41	24.7	144	2	Q54603	Q54603 neisseria m	617	41	24.7	435	16	Q8XRJ7	Q8XRJ7 ralsconia s
545	41	24.7	145	11	Q9CQ83	Q9CQ83 mus musculu	618	41	24.7	439	16	Q97J30	Q97J30 ciostriidum
546	41	24.7	151	5	Q9WIC4	Q9WIC4 drosophila	619	41	24.7	440	2	Q931M6	Q931M6 nostoc punc
547	41	24.7	152	16	Q9ZMH5	Q9ZMH5 helicobacte	620	41	24.7	441	10	Q9SLB4	Q9SLB4 arbidopsis
548	41	24.7	173	2	Q9E8S1	Q9E8S1 uncultured	621	41	24.7	447	17	Q8TID7	Q8TID7 methanosarc
549	41	24.7	177	16	Q9S1R9	Q9S1R9 streptococ	622	41	24.7	452	5	Q23660	Q23660 caenorhadi
550	41	24.7	179	16	P74225	P74225 synchocyst	623	41	24.7	459	16	Q98127	Q98127 rhizobium 1
551	41	24.7	183	17	P74225	P74225 methobact	624	41	24.7	461	13	Q9DGJ4	Q9DGJ4 brachydanio
552	41	24.7	190	2	Q933S1	Q933S1 escherichia	625	41	24.7	462	4	Q14742	Q14742 homo sapien
553	41	24.7	196	16	Q8XJ70	Q8XJ70 ciostriidum	626	41	24.7	462	11	Q35255	Q35255 mus musculu
554	41	24.7	197	4	Q8TD76	Q8TD76 homo sapien	627	41	24.7	462	11	Q951L9	Q951L9 mus musculu
555	41	24.7	201	16	Q981K6	Q981K6 rhizobium 1	628	41	24.7	463	16	Q8UBM6	Q8UBM6 agrobacteri
556	41	24.7	205	16	Q8E5D5	Q8E5D5 salmoneila	629	41	24.7	467	10	Q9AX37	Q9AX37 cryza sativ
557	41	24.7	206	4	Q16354	Q16354 homo sapien	630	41	24.7	468	16	Q9BA92	Q9BA92 rhizobium 1
558	41	24.7	207	12	Q9YUD7	Q9YUD7 alternanthe	631	41	24.7	476	16	Q9HU59	Q9HU59 pseudomonas
559	41	24.7	212	2	Q91880	Q91880 pseudomonas	632	41	24.7	477	5	Q9TZG7	Q9TZG7 caenorhadi
560	41	24.7	215	12	Q8Y001	Q8Y001 papaya mosa	633	41	24.7	478	10	Q8W3P8	Q8W3P8 phaseolus a
561	41	24.7	215	12	Q8Y0P8	Q8Y0P8 papaya mosa	634	41	24.7	479	10	Q9LS07	Q9LS07 arbidopsis
562	41	24.7	217	4	Q8TD75	Q8TD75 homo sapien	635	41	24.7	483	2	Q9Z681	Q9Z681 mycoplasma
563	41	24.7	224	10	Q22411	Q22411 zea mays (m	636	41	24.7	487	2	Q918R7	Q918R7 mycobacteri
564	41	24.7	228	16	Q9K1N3	Q9K1N3 neisseria m	637	41	24.7	491	1	Q9KXW5	Q9KXW5 pseudomonas
565	41	24.7	234	10	Q8S7U8	Q8S7U8 oryza sativ	638	41	24.7	495	10	Q81746	Q81746 arbidopsis
566	41	24.7	234	16	Q9U056	Q9U056 staphylococ	639	41	24.7	496	8	Q9P5T8	Q9P5T8 neurospora
567	41	24.7	241	16	Q9PMW2	Q9PMW2 campylobact	640	41	24.7	498	8	Q8W1Q6	Q8W1Q6 disteligma
568	41	24.7	254	10	Q9S143	Q9S143 arbidopsis	641	41	24.7	502	6	Q95JW3	Q95JW3 macaca fasc
569	41	24.7	259	5	Q9X161	Q9X161 ctenocephal	642	41	24.7	504	17	Q9HMK4	Q9HMK4 halobacteri
570	41	24.7	268	4	Q8TD78	Q8TD78 homo sapien	643	41	24.7	507	16	Q981U5	Q981U5 rhizobium 1
571	41	24.7	273	5	Q9U0P7	Q9U0P7 drosophila	644	41	24.7	508	11	Q91WR4	Q91WR4 mus musculu
572	41	24.7	283	2	Q91865	Q91865 azospirilla	645	41	24.7	509	11	Q922S1	Q922S1 mus musculu
573	41	24.7	288	4	Q96P36	Q96P36 homo sapien	646	41	24.7	509	8	Q8GHC0	Q8GHC0 metanartrec
574	41	24.7	293	11	Q54976	Q54976 mus musculu	647	41	24.7	511	4	Q9UGK5	Q9UGK5 homo sapien
575	41	24.7	293	11	Q8YDZ5	Q8YDZ5 mus musculu	648	41	24.7	511	4	Q96QD5	Q96QD5 homo sapien
576	41	24.7	295	16	Q8XKGO	Q8XKGO salmoneila	649	41	24.7	513	8	Q9WV07	Q9WV07 brachyelytr
577	41	24.7	297	17	Q8Z1W7	Q8Z1W7 pyrobaculum	650	41	24.7	514	16	Q8YCU6	Q8YCU6 homo sapien
578	41	24.7	298	10	Q945K3	Q945K3 zea mays (m	651	41	24.7	514	16	Q8ZOM0	Q8ZOM0 salmoneila
579	41	24.7	301	16	Q9PB12	Q9PB12 xyliella fas	652	41	24.7	530	16	Q8Z870	Q8Z870 salmoneila
580	41	24.7	302	11	Q8YCB8	Q8YCB8 mus musculu	653	41	24.7	531	16	Q8Z870	Q8Z870 salmoneila
581	41	24.7	303	10	Q9S7C0	Q9S7C0 arbidopsis	654	41	24.7	543	10	Q9FYR6	Q9FYR6 arbidopsis
582	41	24.7	306	10	Q81JBS	Q81JBS arbidopsis	655	41	24.7	548	15	Q9QB4	Q9QB4 similan foam
583	41	24.7	310	10	Q8S1B9	Q8S1B9 oryza sativ	656	41	24.7	589	10	Q9M9Y5	Q9M9Y5 arbidopsis
584	41	24.7	317	16	Q98M28	Q98M28 rhizobium 1	657	41	24.7	632	12	Q9E237	Q9E237 heliocoverpa
585	41	24.7	317	16	Q98M28	Q98M28 rhizobium 1	658	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
586	41	24.7	317	16	Q98M28	Q98M28 rhizobium 1	659	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
587	41	24.7	324	8	Q8SP56	Q8SP56 shigella so	660	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
588	41	24.7	324	8	Q8SP56	Q8SP56 shigella so	661	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
589	41	24.7	327	10	Q86641	Q86641 arbidopsis	662	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
590	41	24.7	330	16	Q9CQ52	Q9CQ52 mycobacteri	663	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
591	41	24.7	330	16	Q9CQ52	Q9CQ52 mycobacteri	664	41	24.7	636	2	Q9XAZ9	Q9XAZ9 heliocoverpa
592	41	24.7	339	4	Q9S0E2	Q9S0E2 homo sapien	665	41	24.7	645	4	Q9H470	Q9H470 homo sapien
593	41	24.7	340	16	Q9KSH0	Q9KSH0 vibrrio chol	666	41	24.7	645	4	Q9H470	Q9H470 homo sapien
594	41	24.7	348	5	Q9VBN4	Q9VBN4 drosophila	667	41	24.7	649	11	Q9ES07	Q9ES07 rattus norv
595	41	24.7	349	4	Q9UBJ5	Q9UBJ5 homo sapien	668	41	24.7	669	17	Q9S474	Q9S474 pyrococcus
596	41	24.7	349	10	Q81414	Q81414 arbidopsis	669	41	24.7	670	10	Q9SWO0	Q9SWO0 oryza sativ
597	41	24.7	352	4	Q8TRB34	Q8TRB34 homo sapien	670	41	24.7	671	16	Q9CHU4	Q9CHU4 lactococcus
598	41	24.7	353	2	Q9F5A5	Q9F5A5 agrobacteri	671	41	24.7	672	12	Q99GV3	Q99GV3 heliocoverpa
599	41	24.7	360	3	Q9C228	Q9C228 neurospora	672	41	24.7	672	12	Q91BW4	Q91BW4 heliocoverpa
600	41	24.7	369	16	Q8ZAS8	Q8ZAS8 versinia pe	673	41	24.7	703	10	Q8V5O9	Q8V5O9 helicoverpa
			376	4	Q96P35	Q96P35 homo sapien	673	41	24.7	703	10	Q9LID3	Q9LID3 arbidopsis

674	41	24.7	715	5	Q94494	Q94494 dictyosteli	747	40.5	24.4	406	5	Q9N419	Q9N419 caenorhabdi
675	41	24.7	731	10	Q8W220	Q8W220 oryza sativ	748	40.5	24.4	410	8	Q9MJ82	Q9MJ82 lingua ung
676	41	24.7	736	2	Q93PC0	Q93PC0 microscilla	749	40.5	24.4	417	8	Q9TAX2	Q9TAX2 plasmodium
677	41	24.7	773	2	Q93FC6	Q93FC6 acaroacus ev	750	40.5	24.4	417	17	Q8RZM4	Q8RZM4 pyrobaculum
678	41	24.7	884	3	Q94685	Q94685 schizosacch	751	40.5	24.4	434	2	Q9LAS6	Q9LAS6 streptomyc
679	41	24.7	925	10	Q9Z7A5	Q9Z7A5 arabidopsis	752	40.5	24.4	435	10	Q04203	Q04203 arabidopsis
680	41	24.7	954	11	Q62141	Q62141 mus musculu	753	40.5	24.4	436	16	Q93J79	Q93J79 streptomyc
681	41	24.7	985	16	Q98207	Q98207 rhizobium 1	754	40.5	24.4	437	8	Q47461	Q47461 prococentru
682	41	24.7	990	10	Q06599	Q06599 oryza sativ	755	40.5	24.4	442	3	Q9C1K8	Q9C1K8 neopospora
683	41	24.7	1079	5	Q17566	Q17566 caenorhabdi	756	40.5	24.4	457	10	Q65624	Q65624 arabidopsis
684	41	24.7	1114	10	Q8S6D2	Q8S6D2 oryza sativ	757	40.5	24.4	467	10	Q949P1	Q949P1 arabidopsis
685	41	24.7	1130	4	Q75182	Q75182 homo sapien	758	40.5	24.4	477	5	Q9B140	Q9B140 caenorhabdi
686	41	24.7	1138	12	Q996P9	Q996P9 culex nigrit	759	40.5	24.4	477	5	Q62618	Q62618 plasmodium
687	41	24.7	1172	10	Q9LWY9	Q9LWY9 oryza sativ	760	40.5	24.4	477	8	Q63695	Q63695 plasmodium
688	41	24.7	1306	10	Q9FGF4	Q9FGF4 arabidopsis	761	40.5	24.4	480	8	Q9B970	Q9B970 carotoseolen
689	41	24.7	1403	16	Q97HX7	Q97HX7 clostridium	762	40.5	24.4	484	16	Q9CKQ2	Q9CKQ2 pasteurella
690	41	24.7	1415	4	Q9H7M7	Q9H7M7 homo sapien	763	40.5	24.4	485	15	Q67935	Q67935 aquifex aeo
691	41	24.7	1449	2	Q9P3Z5	Q9P3Z5 neisseria m	764	40.5	24.4	486	16	Q99W44	Q99W44 staphylococ
692	41	24.7	1459	16	Q9JW54	Q9JW54 neisseria m	765	40.5	24.4	505	2	Q9RBR2	Q9RBR2 pseudomonas
693	41	24.7	1463	16	Q98XN1	Q98XN1 rhizobium 1	766	40.5	24.4	507	8	Q47127	Q47127 pieris nana
694	41	24.7	1488	4	Q9HCM3	Q9HCM3 homo sapien	767	40.5	24.4	508	8	Q9G6Y0	Q9G6Y0 plasmodium
695	41	24.7	1491	11	Q8V143	Q8V143 rattus norv	768	40.5	24.4	510	8	Q8WH10	Q8WH10 wolffella
696	41	24.7	1593	11	Q9FP48	Q9FP48 oryza sativ	769	40.5	24.4	510	8	Q8WHK9	Q8WHK9 wolffella
697	41	24.7	1952	5	Q95SNN	Q95SNN drosophila	770	40.5	24.4	510	8	Q8WHK7	Q8WHK7 wolffella
698	41	24.7	2045	5	Q9VRN7	Q9VRN7 drosophila	771	40.5	24.4	510	8	Q8WHK6	Q8WHK6 wolffella
699	41	24.7	2209	5	Q973Z4	Q973Z4 plasmodium	772	40.5	24.4	510	8	Q8WHK4	Q8WHK4 wolffella
700	41	24.7	3096	5	Q9U0U7	Q9U0U7 leishmania	773	40.5	24.4	510	8	Q8WHK2	Q8WHK2 wolffella mic
701	41	24.7	10917	2	Q93NM6	Q93NM6 streptomyc	774	40.5	24.4	510	8	Q8WHK0	Q8WHK0 wolffella aus
702	40.5	24.4	105	16	Q98LBS	Q98LBS rhizobium 1	775	40.5	24.4	512	8	Q9MDJ2	Q9MDJ2 plasmodium
703	40.5	24.4	116	16	Q98NN8	Q98NN8 rhizobium 1	776	40.5	24.4	512	8	Q9MTJ9	Q9MTJ9 plasmodium
704	40.5	24.4	140	8	Q9XKM3	Q9XKM3 trichinella	777	40.5	24.4	512	8	Q8WHK8	Q8WHK8 plasmodium
705	40.5	24.4	140	8	Q9XKM3	Q9XKM3 trichinella	778	40.5	24.4	512	8	Q8WHK8	Q8WHK8 wolffella
706	40.5	24.4	140	8	Q9XKM4	Q9XKM4 trichinella	779	40.5	24.4	513	5	Q9V8T4	Q9V8T4 wolffella bor
707	40.5	24.4	140	8	Q9XKM5	Q9XKM5 trichinella	780	40.5	24.4	514	8	Q9B8A3	Q9B8A3 drosophila
708	40.5	24.4	140	8	Q9XKM6	Q9XKM6 trichinella	781	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
709	40.5	24.4	140	8	Q9XKM7	Q9XKM7 trichinella	782	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
710	40.5	24.4	140	8	Q9XKM8	Q9XKM8 trichinella	783	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
711	40.5	24.4	140	8	Q9XKM9	Q9XKM9 trichinella	784	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
712	40.5	24.4	140	8	Q9XKM9	Q9XKM9 trichinella	785	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
713	40.5	24.4	185	11	Q9Z0S9	Q9Z0S9 mus musculu	786	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
714	40.5	24.4	189	8	Q9MKP2	Q9MKP2 synalpheus	787	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
715	40.5	24.4	219	16	Q8YOK1	Q8YOK1 ralsconia s	788	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
716	40.5	24.4	226	16	Q9LUD3	Q9LUD3 keteleeria	789	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
717	40.5	24.4	242	16	Q8ZPE7	Q8ZPE7 salmonella	790	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
718	40.5	24.4	242	16	Q8ZU76	Q8ZU76 salmonella	791	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
719	40.5	24.4	252	16	Q8YUW3	Q8YUW3 anabaena sp	792	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
720	40.5	24.4	269	5	P90823	P90823 caenorhabdi	793	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
721	40.5	24.4	270	17	Q58473	Q58473 methanococ	794	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
722	40.5	24.4	275	8	Q21352	Q21352 euhadra her	795	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
723	40.5	24.4	277	16	Q98652	Q98652 rhizobium 1	796	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
724	40.5	24.4	280	16	Q98652	Q98652 rhizobium 1	797	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
725	40.5	24.4	286	5	Q19902	Q19902 caenorhabdi	798	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
726	40.5	24.4	294	17	Q8TP15	Q8TP15 methanosarc	799	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
727	40.5	24.4	301	8	Q8WEX1	Q8WEX1 mytilus gal	800	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
728	40.5	24.4	301	8	Q8WEX1	Q8WEX1 keeneana sam	801	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
729	40.5	24.4	301	16	Q911E1	Q911E1 pseudomonas	802	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
730	40.5	24.4	301	16	Q8XZ02	Q8XZ02 ralsconia s	803	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
731	40.5	24.4	325	17	Q8TWV6	Q8TWV6 methanopyru	804	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
732	40.5	24.4	332	10	Q9ST47	Q9ST47 zea mays (m	805	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
733	40.5	24.4	333	2	Q54144	Q54144 streptomyc	806	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
734	40.5	24.4	350	5	Q77109	Q77109 anagrapta f	807	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
735	40.5	24.4	355	16	Q9K1T4	Q9K1T4 vibrio chol	808	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
736	40.5	24.4	362	16	Q9CNG0	Q9CNG0 pasteurella	809	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
737	40.5	24.4	376	17	Q9HMG8	Q9HMG8 halobacteri	810	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
738	40.5	24.4	377	16	Q98C81	Q98C81 rhizobium 1	811	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
739	40.5	24.4	389	16	Q912K1	Q912K1 pseudomonas	812	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
740	40.5	24.4	397	11	Q70348	Q70348 mus musculu	813	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
741	40.5	24.4	397	11	Q99KJ8	Q99KJ8 mus musculu	814	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
742	40.5	24.4	399	10	Q9ZDQ3	Q9ZDQ3 arabidopsis	815	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
743	40.5	24.4	399	10	Q8YDQ3	Q8YDQ3 anabaena sp	816	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
744	40.5	24.4	400	10	Q9AR57	Q9AR57 solanum tub	817	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
745	40.5	24.4	400	10	Q9AR55	Q9AR55 solanum tub	818	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium
746	40.5	24.4	400	10	Q9AR54	Q9AR54 solanum tub	819	40.5	24.4	514	8	Q914X7	Q914X7 plasmodium

820	40	24.1	168	092XM2	092xm2 rhizobium m	893	40	24.1	427	16	09USU7	09jsu7 neisseria m
821	40	24.1	109	096362	096362 hyphantria	894	40	24.1	432	17	026751	026751 methanobact
822	40	24.1	109	0956v5	0956v5 prococcitru	895	40	24.1	438	17	097UX2	097ux2 sulfobus
823	40	24.1	125	095U90	095u90 leishmania	896	40	24.1	445	12	091IH9	091ih9 frog adenov
824	40	24.1	135	086419	086419 nostoc sp.	897	40	24.1	448	16	092IG1	092ig1 rickettsia
825	40	24.1	139	046773	046773 escherichia	898	40	24.1	460	10	098RT0	098rt0 arabidopsis
826	40	24.1	142	091AW6	091aw6 erwinia amy	899	40	24.1	468	16	0927C3	0927c3 listeria mo
827	40	24.1	145	09BRV6	09brv6 homo sapien	900	40	24.1	470	5	005351	005351 entamoeba h
828	40	24.1	154	090T41	090t41 cricetidae	901	40	24.1	475	8	047131	047131 satyria var
829	40	24.1	160	09AKT1	09akt1 paenibacill	902	40	24.1	476	13	013256	013256 gallus gall
830	40	24.1	166	09VR36	09vr36 deinococcus	903	40	24.1	477	16	08X9K8	08x9k8 escherichia
831	40	24.1	185	09VPR1	09vpr1 drosophila	904	40	24.1	481	6	097641	097641 equus caball
832	40	24.1	189	021530	021530 caenorhabd1	905	40	24.1	484	10	082174	082174 arabidopsis
833	40	24.1	189	08RG45	08rg45 fusobacteri	906	40	24.1	485	16	006140	006140 mycobacteri
834	40	24.1	192	000503	000503 calonectria	907	40	24.1	490	17	08TIR5	08tir5 methanosarc
835	40	24.1	198	096516	096516 caenorhabd1	908	40	24.1	491	10	094EV5	094ev5 zea mays (m
836	40	24.1	200	096516	096516 lactococcus	909	40	24.1	494	17	0974C6	0974c6 sulfobus
837	40	24.1	202	09CWM6	09cwm6 drosophila	910	40	24.1	495	2	057340	057340 pantoea ag
838	40	24.1	205	096NU9	096nu9 rhizobium l	911	40	24.1	495	5	09VA60	09va60 drosophila
839	40	24.1	212	08YMM7	08ymm7 anabaena sp	912	40	24.1	497	16	082RZ7	082rz7 salmonella
840	40	24.1	214	036189	036189 alphus can	913	40	24.1	497	16	0829N9	0829n9 salmonella
841	40	24.1	222	09HMC5	09hmc5 halobacteri	914	40	24.1	498	11	08VCO9	08vc09 mus musculu
842	40	24.1	225	098EF5	098ef5 rhizobium l	915	40	24.1	499	16	08RIE8	08rie8 fusobacteri
843	40	24.1	232	09ND98	09nd98 drosophila	916	40	24.1	501	16	09ZBY7	09zby7 streptomyc
844	40	24.1	232	0917L4	0917l4 drosophila	917	40	24.1	504	8	047141	047141 actinidia c
845	40	24.1	241	092Q58	092q58 rhizobium m	918	40	24.1	504	8	095C11	095c11 actinidia c
846	40	24.1	243	092VK6	092vk6 rhizobium m	919	40	24.1	504	8	095C08	095c08 actinidia h
847	40	24.1	244	093J05	093j05 pseudomonas	920	40	24.1	504	8	095C06	095c06 actinidia v
848	40	24.1	245	08X2W6	08x2w6 escherichia	921	40	24.1	504	8	095C07	095c07 actinidia p
849	40	24.1	253	08TNV6	08tnv6 anabaena sp	922	40	24.1	504	8	095C06	095c06 actinidia c
850	40	24.1	258	09RZ31	09rz31 deinococcus	923	40	24.1	504	8	095C05	095c05 actinidia i
851	40	24.1	272	066627	066627 aquifex aeo	924	40	24.1	504	8	095C04	095c04 actinidia f
852	40	24.1	275	063841	063841 mus sp. udp	925	40	24.1	504	8	095C03	095c03 actinidia s
853	40	24.1	278	0913H1	0913h1 pseudomonas	926	40	24.1	504	8	095C02	095c02 actinidia l
854	40	24.1	282	08ZRH0	08zrh0 yersinia pe	927	40	24.1	504	8	095C01	095c01 actinidia c
855	40	24.1	285	08XG59	08xg59 salmonella	928	40	24.1	504	8	095C00	095c00 actinidia c
856	40	24.1	286	08ZC21	08zc21 yersinia pe	929	40	24.1	504	8	095B24	095b24 actinidia e
857	40	24.1	290	098NT4	098nt4 rhizobium l	930	40	24.1	504	8	095B23	095b23 actinidia r
858	40	24.1	293	09VJK8	09vjk8 drosophila	931	40	24.1	504	8	094Q39	094q39 actinidia z
859	40	24.1	296	09A665	09a665 caulobacter	932	40	24.1	504	8	094P88	094p88 actinidia m
860	40	24.1	299	09KEP9	09kep9 bacillus ha	933	40	24.1	504	8	094N79	094n79 actinidia m
861	40	24.1	301	08ZU75	08zu75 pyrobaculum	934	40	24.1	504	8	094N79	094n79 actinidia m
862	40	24.1	301	08TWM3	08twm3 methanopyru	935	40	24.1	504	10	08W0S3	08w0s3 sorghum bic
863	40	24.1	314	09HAD7	09had7 homo sapien	936	40	24.1	506	8	095C10	095c10 actinidia k
864	40	24.1	317	08V1J7	08v1j7 mus musculu	937	40	24.1	506	8	095B28	095b28 clematoclet
865	40	24.1	322	08YY66	08yy66 anabaena sp	938	40	24.1	508	13	09DEFS	09def5 gallus gall
866	40	24.1	323	09GVP5	09gvp5 trichomonas	939	40	24.1	509	16	09KSC6	09ksc6 vibrio chol
867	40	24.1	329	09U1B4	09u1b4 bacillus ha	940	40	24.1	511	2	09EV93	09ev93 clostridium
868	40	24.1	339	09KE30	09ke30 aradidopsis	941	40	24.1	511	8	09WV02	09wv02 chusquea co
869	40	24.1	330	09W2J5	09w2j5 trichomonas	942	40	24.1	511	8	09WV01	09wv01 phyllostach
870	40	24.1	333	096518	096518 homo sapien	943	40	24.1	511	8	09WV00	09wv00 sasa kurile
871	40	24.1	334	09H7H8	09h7h8 homo sapien	944	40	24.1	511	8	09WU27	09wu27 nardus stri
872	40	24.1	336	058151	058151 pyrococcus	945	40	24.1	511	8	09WU22	09wu22 melica alti
873	40	24.1	333	034885	034885 bacillus su	946	40	24.1	511	8	09WU21	09wu21 brachypodu
874	40	24.1	347	096FW0	096fw0 homo sapien	947	40	24.1	511	8	09WUY5	09wuy5 nassella te
875	40	24.1	347	091201	091201 caenorhabd1	948	40	24.1	511	8	09WUY4	09wuy4 stipa ofine
876	40	24.1	370	09X7X1	09x7x1 streptomyce	949	40	24.1	512	8	09WUW7	09wuw7 microchloa
877	40	24.1	373	09SDW0	09sdw0 xyella fas	950	40	24.1	513	2	09KWL8	09kw18 corynebacte
878	40	24.1	382	09N3M5	09n3m5 caenorhabd1	951	40	24.1	513	8	09T382	09t382 eustachys m
879	40	24.1	383	091273	091273 caenorhabd1	952	40	24.1	513	8	09TIC1	09tic1 phragmites
880	40	24.1	388	09POF1	09pof1 ureaplasma	953	40	24.1	513	8	09TIC0	09tic0 zeugites pi
881	40	24.1	400	09A7Z0	09a7z0 caulobacter	954	40	24.1	513	8	09TIB9	09tib9 andropogon
882	40	24.1	402	09AR56	09ar56 solanum tub	955	40	24.1	513	8	09TIB7	09tib7 pleuraphis
883	40	24.1	402	08TW29	08tw29 methanopyru	956	40	24.1	513	8	09TIB6	09tib6 eleusine in
884	40	24.1	404	09HRV5	09hrv5 halobacteri	957	40	24.1	513	8	09TIB5	09tib5 coelachyrum
885	40	24.1	418	09WYCV	09wyvc thermotoga	958	40	24.1	513	8	09TIB4	09tib4 chloris mos
886	40	24.1	420	040108	040108 lupinus ang	959	40	24.1	513	8	09TIB3	09tib3 cynodon dac
887	40	24.1	420	040107	040107 lupinus ang	960	40	24.1	513	8	09TIB1	09tib1 eustachys d
888	40	24.1	420	006832	006832 mycobacteri	961	40	24.1	513	8	09TIB0	09tib0 trichorhis
889	40	24.1	424	09EXY03	09exy3 clostridium	962	40	24.1	513	8	09TIB8	09tib8 astrubedia la
890	40	24.1	426	005907	005907 mycobacteri	963	40	24.1	513	8	09TIB8	09tib8 perotus rar
891	40	24.1	437	P95363	P95363 neisseria g	964	40	24.1	513	8	09TIA7	09tia7 tiragus bert
892	40	24.1	427	09K133	09k133 neisseria m	965	40	24.1	513	8	09TIA3	09tia3 trichoneura

```

966 40 24.1 513 8 09T199 09T199 orcuttia ca
967 40 24.1 513 8 09T197 09T197 sporobolus
968 40 24.1 513 8 09T196 09T196 monodia sti
969 40 24.1 513 8 09T195 09T195 plectrachne
970 40 24.1 513 8 09T194 09T194 pappophorum
971 40 24.1 513 8 09T193 09T193 eragrostis
972 40 24.1 513 8 09T192 09T192 fingerhuthi
973 40 24.1 513 8 09T191 09T191 uniola pani
974 40 24.1 513 8 09MUY3 09MUY3 arundo dona
975 40 24.1 513 8 09MUY2 09MUY2 denthonia s
976 40 24.1 513 8 09MUY1 09MUY1 centropodia
977 40 24.1 513 8 09MUY0 09MUY0 molinia cae
978 40 24.1 513 8 09MUY7 09MUY7 chasmanthiu
979 40 24.1 513 8 09MUY6 09MUY6 lophatherum
980 40 24.1 513 8 09MUY5 09MUY5 orthoclada
981 40 24.1 513 8 09MUY1 09MUY1 tristachya
982 40 24.1 513 8 09MUY0 09MUY0 digitaria s
983 40 24.1 513 8 09MUY8 09MUY8 panicum cap
984 40 24.1 513 8 09MUY6 09MUY6 zoysia macr
985 40 24.1 513 8 09MUY5 09MUY5 eragrostis
986 40 24.1 513 8 09MUY4 09MUY4 kengia song
987 40 24.1 513 8 09MUY3 09MUY3 sporobolus
988 40 24.1 513 8 09T1A5 09T1A5 eragrostis
989 40 24.1 513 8 09T1A6 09T1A6 eragrostis
990 40 24.1 513 8 09T198 09T198 muldenbergi
991 40 24.1 513 8 09T1A2 09T1A2 tridens bra
992 40 24.1 513 8 09T1A1 09T1A1 aegopogon c
993 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
994 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
995 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
996 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
997 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
998 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
999 40 24.1 513 8 09T1A2 09T1A2 aegopogon c
1000 40 24.1 513 8 09T1A2 09T1A2 aegopogon c

```

ALIGNMENTS

```

RESULT 1
ID 09Y5R3 PRELIMINARY; PRT: 386 AA.
AC 09Y5R3:
AD 09Y5R3:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand
DE sulfotransferase GST-3).
GN GST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RX MEDLINE=99764336; PubMed=10330415;
RA Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R.,
RA Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.;
RA "Sulfotransferases of two specificities function in the reconstitution
RT of high endothelial cell ligands for L-selectin.";
RT J. Cell Biol. 145:899-910(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
RA Rosen S.D.;
RA "Chromosomal Localization and Genomic Organization for the
RT Galactose/N-Acetylglucosamine 6-O-sulfotransferase Gene Family.";
RT Sulfotransferase Gene Family.";
RT Glycobiology 0:0-0(2001).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=11332592; PubMed=11439191;
RA Yeh J.-C., Hiraoka N., Petryniak B., Nakayama J., Ellies L.G.,
RA Raduka D., Hindsgrall O., March J.D., Lowe J.B., Raduka M.;
RT "Novel sulfated lymphocyte homing receptors and their control by a
RT corel extension betaf1,3-N-acetylglucosaminyltransferase.";
RL Cell 105:957-969(2001).
DR EMBL: AF131235; AAD3015.1; -.
DR EMBL: AF280088; AAG48246.1; -.
DR EMBL: AF149783; AAK48417.1; -.
DR INTERPRO: IPR001092; ELH_Dbsic.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
KW Lectin; Selectin; Transferase.
SQ SEQUENCE 386 AA; 45133 MW; 0C3BB4022417143A CRC64;

```

```

Query Match 100.0%; Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SMRSGSFVQGLFGQHPDYFLMEPAHV 29
DB 50 SMRSGSFVQGLFGQHPDYFLMEPAHV 78

```

```

RESULT 2
ID 09R1I1 PRELIMINARY; PRT: 388 AA.
AC 09R1I1:
AD 09R1I1:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE L-selectin ligand sulfotransferase.
DE CHST4.
GN CHST4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9361934; PubMed=10435581;
RA Hiraoka N., Petryniak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,
RA Izawa D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.;
RT "A novel, high endothelial venule-specific sulfotransferase expresses
RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.";
RL Immunity 11:79-89(1999).
DR EMBL: AF109155; AAD45579.1; -.
DR MGI: MGI:1349479; Chst4.
KW Lectin; Selectin; Transferase.
SQ SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884AE CRC64;

```

```

Query Match 100.0%; Score 166; DB 11; Length 388;
Best Local Similarity 100.0%; Pred. No. 2.7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SMRSGSFVQGLFGQHPDYFLMEPAHV 29
DB 49 SMRSGSFVQGLFGQHPDYFLMEPAHV 77

```

```

RESULT 3
ID 09W0E5 PRELIMINARY; PRT: 388 AA.
AC 09W0E5:
AD 09W0E5:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase (Chondroitin
DE 6/keratan) sulfotransferase 4).
GN CHST4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```


RP SEQUENCE FROM N.A.
 RX MEDLINE=20472330; PubMed=11017086;
 RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K.,
 RA Nakamura T., Dots A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S.,
 RA Fujikawa T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A.,
 RA Fukuda M.N.;
 RT "Macular corneal dystrophy type I and type II are caused by distinct
 RT mutations in a new sulphotransferase gene";
 RL Nat. Genet. 26:237-241(2000).
 RN [2]
 RP SEQUENCE OF 22-411 FROM N.A.
 RC TISSUE=TESTIS;
 RX MEDLINE=942349; PubMed=10491328;
 RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
 RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
 RT Sulfotransferase that is Highly Restricted to Intestinal Tissue";
 RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
 DR EMBL: AF246718; AAG28023.1; -;
 DR EMBL: AF219991; AAG26326.1; -;
 DR EMBL: AF176839; AAD56001.1; -;
 DR EMBL: AF176838; AAD56000.1; -;
 KM Transferase.
 SQ SEQUENCE 411 AA; 46160 MW; 97642D54BE26E06 CRC64;
 Query Match 94.6%; Score 157; DB 4; Length 411;
 Best Local Similarity 93.1%; Pred. No. 5.4e-14;
 Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SMRSGSFVQGLFGQHPDVFYLMPEAMHV 29
 Db 70 SMRSGSFVQGLFGQHPDVFYLMPEAMHV 98
 RESULT 7
 Q9EP78 PRELIMINARY; PRT; 484 AA.
 AC Q9EP78;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE N-acetylglucosamine 6-O-sulfotransferase GST-5 (N-
 DE acetylglucosamine-6-O-sulfotransferase) (2600013M07R1k protein)
 DE (Carbohydrate (N-acetylglucosamine) sulfotransferase 7).
 GN CHST7 OR GST5 OR 2600013M07R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX PubMed=10956661;
 RA Bhakta S., Bartsch A., Bowman K.G., Kao W.M., Polsky I., Lee J.-K.,
 RA Cook B.N., Buehl R.E., Rosen S.D., Bertozzi C.R., Hemmerich S.;
 RT "Sulfation of N-Acetylglucosamine by Chondroitin 6-Sulfotransferase 2
 RT (GST-5).";
 RL J. Biol. Chem. 275:40226-40234(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Uchimura K., Muramatsu T.;
 RT "Identification and molecular characterization of a cDNA encoding a
 RT novel N-acetylglucosamine-6-O-sulfotransferase";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Traub F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Oikido T., Furuno M., Bono H., Baldarelli P., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bernaldo M.F.,
 RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF280089; AAG48247.1; -;
 DR EMBL: AB040710; BAB13769.1; -;
 DR EMBL: AK011202; BAB27465.1; -;
 DR EMBL: BC019204; AAH19204.1; -;
 DR MGD: MGI:1891767; Chst7.
 KM Transferase.
 SQ SEQUENCE 484 AA; 54766 MW; 9B195537D7AB7193 CRC64;
 Query Match 80.7%; Score 134; DB 11; Length 484;
 Best Local Similarity 72.4%; Pred. No. 1.1e-10;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SMRSGSFVQGLFGQHPDVFYLMPEAMHV 29
 Db 107 TWRGSSFLGELFQHPDVFYLMPEAMHL 135
 RESULT 8
 Q9NS84 PRELIMINARY; PRT; 486 AA.
 AC Q9NS84;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Chondroitin 6-sulfotransferase-2.
 GN ChST-2.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20347214; PubMed=10781596;
 RA Kitagawa H., Fujita M., Ito N., Sugahara K.;
 RT "Molecular cloning and expression of a novel chondroitin 6-O-
 RT sulfotransferase.";
 RL J. Biol. Chem. 275:21075-21080(2000).
 DR EMBL: AB037187; BAB03217.1; -;
 KM Transferase.
 SQ SEQUENCE 486 AA; 54294 MW; B524D1430B3C9075 CRC64;
 Query Match 80.7%; Score 134; DB 4; Length 486;
 Best Local Similarity 72.4%; Pred. No. 1.1e-10;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SMRSGSFVQGLFGQHPDVFYLMPEAMHV 29
 Db 109 TWRGSSFLGELFQHPDVFYLMPEAMHL 137
 RESULT 9
 Q75667 PRELIMINARY; PRT; 486 AA.
 ID Q75667

NAME: _____
CURR: OPTIM: 7:00 (1990):

```

RN [2]
RP SEQUENCE FROM N.A.
RA Peng T., Tabas I., Williams K.J.
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RX MEDLINE=98070405; PubMed=9405439;
RA Fukuta M., Inazawa J., Torii T., Tsuzuki K., Shimada E., Habuchi O.;
RT "Molecular cloning and characterization of human keratan sulfate
  Gal-6-sulfotransferase."
RL J. Biol. Chem. 272:32321-32328(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-EMBILICAL VEIN ENDOTHELIUM;
RX MEDLINE=99168906; PubMed=10049591;
RA Li X., Tedder T.F.;
RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
  endothelial cells: cDNA cloning, expression, and chromosomal
  localization."
RL Genomics 55:345-347(1999).
DR EMBL: U65637; AAC28776.1;
DR EMBL: AB003791; BAA24840.1;
DR EMBL: AF090137; AAD19878.1;
KM TRANSFERASE.
SQ SEQUENCE 411 AA; 46714 MW; F07D0A23B6338A09 CRC64;

Query Match
Best Local Similarity 69.9%; Score 116; DB 4; Length 411;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDVFYLMPEAMHY 29
DB 70 RSGSFEVGLFGQHPDVFYLMPEAMHY 29
|||||
ID 093403 PRELIMINARY; PRT; 441 AA.
AC 093403;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN NSIST.
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Hypnosqualia; Plistiorhina; Batoidae;
OC Torpediniformes; Torpedinidae; Torpedo.
OX NCBI_TaxID=7787;
RN [1]
RP SEQUENCE FROM N.A.
RA Nestuk M.A., Davis S., Yancopoulos G.D., Fallon J.R.;
RT "Expression cloning and characterization of NSIST, a novel
  sulfotransferase expressed by a subset of neurons and postsynaptic
  targets."
RL J. Neurosci. 0:0-0(1998).
DR EMBL: AF079675; AAC28491.1;
KM TRANSFERASE.
SQ SEQUENCE 441 AA; 50392 MW; 993B8B58743BA53 CRC64;

Query Match
Best Local Similarity 65.1%; Score 108; DB 13; Length 441;
Matches 18; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDVFYLMPEAMHY 29
DB 107 RTGSSFEVGEFFNQNDIFLYLEPLAMHY 133
|||||
ID 088199 PRELIMINARY; PRT; 472 AA.

```

```

AC 088199;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN CHST3 OR C6ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, AND 129; TISSUE=SPLEEN;
RX MEDLINE=98259835; PubMed=9597547;
RA Uchimura K., Kadomatsu K., Fan O.-W., Muramatsu H., Kurosawa N.,
  Kaname T., Yamamura K., Fukuta M., Habuchi O., Muramatsu T.;
RT "Mouse chondroitin 6-sulfotransferase: molecular cloning,
  characterization and chromosomal mapping."
RL Glycobiology 8:469-496(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638441; PubMed=11696535;
RA Uchimura K., Kadomatsu K., Nishimura H., Muramatsu H., Nakamura E.,
  Kurosawa N., Habuchi O., El-Fasakhly F.M., Yoshikai Y., Muramatsu T.;
RT "Functional analysis of the chondroitin 6-sulfotransferase gene in
  relation to lymphocyte subpopulations, brain development, and
  overexpressed chondroitin sulfates."
RL J. Biol. Chem. 277:1443-1450(2002).
DR EMBL: AB008937; BAA29054.1;
DR EMBL: AB008938; BAA29055.1;
DR EMBL: AB062109; BAB72166.1;
DR MGI:1858224; Chst3.
KM TRANSFERASE.
SQ SEQUENCE 472 AA; 53997 MW; F021147196D9D339 CRC64;

Query Match
Best Local Similarity 60.8%; Score 101; DB 11; Length 472;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDVFYLMPEAMHY 29
DB 136 RTGSSFEVGEFFNQNDIFLYLEPLMHI 162
|||||
ID 090212 PRELIMINARY; PRT; 474 AA.
AC 090212;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT "Chondroitin 6-sulfotransferase of rat sciatic nerve."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Liu J., Li X., Kwok C.-F., Tang S., Shum K.-Y.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF178689; AAD54386.2;

```

KW Transferase.
SQ SEQUENCE 474 AA; 54018 MW; 83EC94C3894CE77E CRC64;

Query Match 60.8%; Score 101; DB 11; Length 474;
Best Local Similarity 59.3%; Pred. No. 5e-06;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSGFVGQLFGQHPDVFYLMPEAMHV 29
138 RTGSGFVGGEFFNQGNIFLYLEPLMWHI 164

RESULT 20

075099 PRELIMINARY; PRT; 479 AA.

AC 075099; PRELIMINARY; PRT; 479 AA.
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE Chondroitin 6-sulfotransferase (EC 2.8.2.17) (C69T).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=FETAL BRAIN;
RX MEDLINE=98390270; PubMed=9714738;
RA Fukuta M., Kobayashi Y., Uchinura K., Kimata K., Hachimi O.;
RT "Molecular cloning and expression of human chondroitin 6-sulfotransferase.";
RL Biochim. Biophys. Acta 1399:57-61(1998).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=99098360; PubMed=9883891;
RA Tsutsuni K., Shimakawa H., Kitagawa H., Sugahara K.;
RT "Functional expression and genomic structure of human chondroitin 6-sulfotransferase.";
RL FEBS Lett. 441:235-241(1998).
CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN SULFATE.
CC -1- CATALYTIC ACTIVITY: 3'-PHOSPHOADENYLYLSULFATE + CHONDROITIN = ADENOSINE 3',5'-BISPHOSPHATE + CHONDROITIN 6'-SULFATE.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: FOUND AT HIGHEST LEVELS IN SKELETAL MUSCLE.
CC ALSO PRESENT AT LOWER LEVELS IN HEART, PLACENTA AND PANCREAS.
DR EMBL: AB012132; BAA32576.1; -
DR EMBL: AB017915; BAA36348.1; -
KW Transferase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 24 38 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (BY SIMILARITY).
FT DOMAIN 39 458 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 387 387 P -> R (IN PLACENTA).
FT VARIANT 443 443 P -> A (IN PLACENTA).
SQ SEQUENCE 479 AA; 54672 MW; 494A58A48E37D570 CRC64;

Query Match 60.8%; Score 101; DB 4; Length 479;
Best Local Similarity 59.3%; Pred. No. 5e-06;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSGFVGQLFGQHPDVFYLMPEAMHV 29
142 RTGSGFVGGEFFNQGNIFLYLEPLMWHI 168

RESULT 21

09YMC3 PRELIMINARY; PRT; 315 AA.

AC 09YMC3; PRELIMINARY; PRT; 315 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG9551 protein.
GN CG9551.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Gloder A., Gong F., Gorelli J.H., Gu Z., Guan P., Harris M.,
RA Harstin D., Houston K.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson K., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton K., Strong R., Sun E.,
RA Stryker R., Teclor C., Turner R., Venler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AB003614; AAF52398.1; -
DR FlyBase; FBgn0031827; CG9551.
SQ SEQUENCE 315 AA; 35608 MW; EA0A82855CC6418 CRC64;

Query Match 42.8%; Score 71; DB 5; Length 315;
Best Local Similarity 52.0%; Pred. No. 0.054;
Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 SMRSGSFVGQLFGQHPDVFYLMPE 25
191 SMRSGSFVLDDILNIPGNFYHPEP 215

RESULT 22
095TN9 PRELIMINARY; PRT; 486 AA.

```

AC Q95TN9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD3051P.
GN CG9551.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stalton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Prounenavong S., Wan K.,
RA Yu C., Lewis S.E., Rudin G.M., Ceiniker S.,
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY058647; AAL13876.1; -
DR Flybase: FBgn0031827; CG9551.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.
SQ SEQUENCE 486 AA; 55539 MW; BEDCC96F096746EB CRC64;

Query Match 42.8%; Score 71; DB 5; Length 486;
Best Local Similarity 52.0%; Pred. No. 0.087;
Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 SWRSSSFYQLEGGHPDYFLMEP 25
DB 191 SWRSSFTLEDILNIPGNFYHEP 215

RESULT 23
Q8TG14 PRELIMINARY; PRT; 46 AA.
AC Q8TG14;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DE 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mating type 1 protein (Fragment).
OS Phaeosphaeria avenaria.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes Incertae sedis;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=55068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 58582;
RA Czembor P.C., Arseniuk E., Bergstrom G.C., Ueng P.P.;
RT "Sequence diversity of mating type genes in Phaeosphaeria avenaria.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY072936; AAL69556.1; -
FT NON_TER 1 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5554 MW; 48A28037E752FF32 CRC64;

Query Match 34.3%; Score 57; DB 3; Length 46;
Best Local Similarity 50.0%; Pred. No. 0.62;
Matches 12; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 6 SFFVGLFGQHPD--VEYLMEPAM 27
DB 20 SFTIGGLMKQEPDKPLMTLTKAM 43

RESULT 24
Q8YWG4 PRELIMINARY; PRT; 1570 AA.
AC Q8YWG4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein A11649.
GN A11649.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuriiz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimizu S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
DNA Res. 8:205-213(2001).
DR EMBL: AP003586; BAB78015.1; -
DR InterPro: IPR001227; Acyltransferase.
DR InterPro: IPR000794; Ketoacyl-synth.
DR InterPro: IPR003880; Pantone-attach.
DR Pfam: PF00698; Acyl-transf. 1.
DR Pfam: PF00109; ketoacyl-synth. 1.
DR Pfam: PF02801; ketoacyl-synth. C. 1.
DR Pfam: PF00550; PP-binding; 1.
DR PROSITE: PS00775; ACP-DOMAIN; 1.
DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1570 AA; 170791 MW; AB778DE7395F6A52 CRC64;

Query Match 34.3%; Score 57; DB 16; Length 1570;
Best Local Similarity 48.5%; Pred. No. 30;
Matches 16; Conservative 2; Mismatches 7; Indels 8; Gaps 2;

QY 4 SGSS-----FVGLFGQHPDYFLMEP--AMW 28
DB 692 SGQSQAQIFVTQTQQGHPDIFYKELPVYKAFH 724

RESULT 25
Q99VX8 PRELIMINARY; PRT; 277 AA.
AC Q99VX8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DE 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Telchoic acid translocation permease protein.
GN TAGG OR SA0594.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iino Y., Ito T., Kanamori M.,
RA Matsumaru H., Matuyama A., Murakami H., Hasegawa A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.T., Kaito C., Sekimizu K.,
RA Hiraoka H., Kihara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003137; BAB41826.1; -
DR InterPro: IPR000412; ABCtransprt2.
DR Pfam: PF01061; ABC2_membrane; 1.
DR Complete proteome.
SQ SEQUENCE 277 AA; 32381 MW; 603C9629F436608 CRC64;

Query Match 31.9%; Score 53; DB 16; Length 277;

```


DR EMBL: AE009866; AAL64063.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 187 AA; 21760 MW; 0A33A9408FE41942 CRC64;

Query Match 31.6%; Score 52.5; DB 17; Length 187;
 Best Local Similarity 37.0%; Pred. No. 12;
 Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

OY 2 WRSSEVFGOLFQHPD---VFYLMPEAW 23
 DB 26 WNTGTSAGLRFGRGAGRPDPFFLL 52

RESULT 30
 O9C409 PRELIMINARY; PRT; 46 AA.

AC 09C409; PRELIMINARY; PRT; 46 AA.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Mating type 1 protein (Fragment).
 OS Phaeosphaeria nodorum (Septoria nodorum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes Incertae sedis;
 OC Phaeosphaeriaceae; Phaeosphaeria.
 OX NCBI_TaxID=13684;
 RN
 RC SEQUENCE FROM N.A.
 RP STRAIN=SN48-1;
 RA Dai Q., Arseniuk E., Cui K., Deng P.P.;
 RT "Genetic segregation and sexuality in Phaeosphaeria nodorum.";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBD databases.
 DR EMBL: AF322008; AAK09389.1; -
 FT NON_TER 1 1
 SO SEQUENCE 46 AA; 5553 MW; 48A678A4439DB4D2 CRC64;

Query Match 31.3%; Score 52; DB 3; Length 46;
 Best Local Similarity 45.8%; Pred. No. 3.1;
 Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

OY 6 SSFVGOLFQHPD--VFYLMPEAW 27
 DB 20 SSYIGOLWEHPDKPLMTLTKAW 43

RESULT 31
 O8TG16 PRELIMINARY; PRT; 46 AA.

AC 08TG16; PRELIMINARY; PRT; 46 AA.
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Mating type 1 protein (Fragment).
 OS Phaeosphaeria nodorum (Septoria nodorum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes Incertae sedis;
 OC Phaeosphaeriaceae; Phaeosphaeria.
 OX NCBI_TaxID=13684;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=SN27-1;
 RA Dai Q., Cui K.-R., Arseniuk E., Deng P.P.;
 RT "Segregation, Aggressiveness, and Sexuality in Phaeosphaeria
 RT nodorum.";
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBD databases.
 DR EMBL: AT072933; AAL69553.1; -
 FT NON_TER 1 1
 SO SEQUENCE 46 AA; 5614 MW; 48A678A45E2B6FA2 CRC64;

Query Match 31.3%; Score 52; DB 3; Length 46;
 Best Local Similarity 45.8%; Pred. No. 3.1;

Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

OY 6 SSFVGOLFQHPD--VFYLMPEAW 27
 DB 20 SSYIGOLWEHPDKPLMTLTKAW 43

RESULT 32

O913V7 PRELIMINARY; PRT; 518 AA.
 AC 0913V7; PRELIMINARY; PRT; 518 AA.
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Hypothetical protein PA1387.
 GN PA1387.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964 (2000).
 DR EMBL: AE004568; AAG04776.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 518 AA; 58158 MW; 3FB7F9C237C3B5B CRC64;

Query Match 31.3%; Score 52; DB 16; Length 518;
 Best Local Similarity 47.8%; Pred. No. 44;
 Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 3 RGSSEVFGOLFQHPDVFYLMPE 25
 DB 267 RGTSHLCRLSLSQPDTYVLANEP 289

RESULT 33
 Q8U1V4 PRELIMINARY; PRT; 614 AA.

AC Q8U1V4; PRELIMINARY; PRT; 614 AA.
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ABC transporter, substrate binding protein.
 GN ATU0187 OR AGR_C_316.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavith T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Seemphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens

```

RT CS8.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2160851; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Curolo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
RA Houttel K., Gordon J., Vaughn M., Iartchouk O., Epp A., Liu F.,
RA William C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Girson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens CS8.";
RL Science 294:2323-2328(2001).
DR EMBL: AF008992; AAL41211.1;
DR EMBL: AE007959; AAK6007.1;
KW Complete proteome.
SQ SEQUENCE 614 AA; 68551 MW; 117D62469C0F3ED5 CRC64;

Query Match
Best Local Similarity 31.3%; Score 52; DB 16; Length 614;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

OY 2 WRSSSFVQGLFGQHPDV 20
DB 33 WRKGISTVGEL--KHDPGF 49

RESULT 34
OY 08Y756 PRELIMINARY; PRT; 641 AA.
ID 08Y756;
AC 08Y756;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical protein A112870.
GN A112870.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCB1_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003591; BAB74569.1;
DR InterPro: IPR003342; PMT.
DR Pfam: PF02366; PMT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 641 AA; 71650 MW; 8A709F53750AF2C CRC64;

Query Match
Best Local Similarity 31.3%; Score 52; DB 16; Length 641;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 WRSSSFVQGLFGQHP 16
DB 298 WRNGESTINSEFTGYH 312

RESULT 35
OY 096X13 PRELIMINARY; PRT; 46 AA.
ID 096X13;
AC 096X13;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

```

```

DE Mating type 1 protein (Fragment).
OS Phaeosphaeria avenaria f. sp. triticae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyrionmycetes incertae sedis;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCB1_TaxID=54790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA1 25-1;
RA Dai Q., Cui K., Arseniuk E., Ueng P.P.;
RT "Mating type gene in Phaeosphaeria avenaria f. sp. triticae.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF558824; AAK51440.1;
FT NON_TER 1 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5645 MW; 48BCB74461858F4C CRC64;

Query Match
Best Local Similarity 30.7%; Score 51; DB 3; Length 46;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

OY 6 SSFVQGLFGQHPD--VFYLMEPAM 27
DB 20 SSYIGQLMKYEPDKPLMTLTKAM 43

RESULT 36
OY 096X15 PRELIMINARY; PRT; 46 AA.
ID 096X15;
AC 096X15;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Mating type 1 protein (Fragment).
GN MAT1.
OS Phaeosphaeria avenaria f. sp. triticae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyrionmycetes incertae Phaeosphaeriaceae;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCB1_TaxID=54790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA1004CND-96;
RA Dai Q., Cui K., Gilbert J., Bergstrom G., Ueng P.P.;
RT "Mating type gene in Phaeosphaeria avenaria f. sp. triticae.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 18596;
RA Czembor P.C., Arseniuk E., Ueng P.P.;
RT "Sequence diversity of mating type genes in Phaeosphaeria avenaria f. sp. triticae.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF558822; AAK51438.1;
DR EMBL: AF072935; AAL69555.1;
DR EMBL: AT077724; AAL82533.1;
FT NON_TER 1 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5644 MW; 48BCB744618B6FA2 CRC64;

Query Match
Best Local Similarity 30.7%; Score 51; DB 3; Length 46;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

OY 6 SSFVQGLFGQHPD--VFYLMEPAM 27
DB 20 SSYIGQLMKYEPDKPLMTLTKAM 43

```



```

RESULT 37
Q9XBL3      PRELIMINARY;      PRT;      237 AA.
ID          Q9XBL3;
AC          09XBL3;
DT          01-NOV-1999 (TREMBlrel. 12, Created)
DT          01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT          01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE          DNA alkylation repair enzyme.
GN          ALKD.
OS          Bacillus cereus.
OC          Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC          Bacillaceae; Bacillus.
OX          NCBI_TaxID=1396;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 10987;
RX          MEDLINE=99231848; PubMed=10217496;
RA          Okstad O.A., Hegna I., Lindaeck T., Rishovd A.L., Kolsto A.B.;
RT          "Genome organisation is not conserved between Bacillus cereus and
RL          Bacillus subtilis."
DR          Microbiology 145:621-631(1999).
SQ          EMBL: AJ010128; CAB40581.1;
DE          SEQUENCE 237 AA; 28172 MW; 132DAF8961498A83 CRC64;

Query Match          30.7%; Score 51; DB 2; Length 237;
Best Local Similarity 38.1%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY          7 SEVGOLFQHPDYFYLMEPAW 27
DB          118 TFLGNIFLQHPDLISATIPKW 138

RESULT 38
Q9L3S3      PRELIMINARY;      PRT;      263 AA.
ID          Q9L3S3;
AC          09L3S3;
DT          01-OCT-2000 (TREMBlrel. 15, Created)
DT          01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT          01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE          Putative ATP-binding subunit.
GN          URTD.
OS          Anabaena sp. (strain PCC 7120).
OC          Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX          NCBI_TaxID=103690;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=PCC 7120;
RA          Valladares A., Montesinos M., Herrero A., Flores F.;
RT          "A procarboxylic urea transporter identified in two cyanobacteria."
RL          Submitted (JAN-2000) to the EMBL/GenBank/DBD databases.
DR          EMBL: AA271599; CAB70951.1;
DR          HSSP: Q58663; 165H;
DR          InterPro: IPR003439; ABC_transport.
DR          Pfam: PF00005; ABC_tran; 1.
KW          ATP-binding.
SQ          SEQUENCE 263 AA; 30320 MW; F3E09CCC95B6A8 CRC64;

Query Match          30.7%; Score 51; DB 2; Length 263;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

OY          1 SWRSGSFVQGLFGHPDYFYLMEP 25
DB          166 SWRKATFRMKLVGQSPDLIVDEP 190

RESULT 39
Q93WX7      PRELIMINARY;      PRT;      341 AA.
AC          Q93WX7;

```

```

DT          01-DEC-2001 (TREMBlrel. 19, Created)
DT          01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT          01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE          Aspartate aminotransferase (Fragment).
GN          CRAT1.
OS          Securigera parviflora.
OC          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC          eurosids I; Fabales; Fabaceae; Papilionoideae; Loreae; Securigera.
OX          NCBI_TaxID=172083;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          TISSUE=LEAF-DERIVED CALLUS;
RA          Arlanti N.;
RT          "Molecular characterization of selected genes in ammonium assimilation
RL          from in vitro cultures of the legume Coronilla rostrata."
DR          Thesis (2001). The University of New South Wales.
DR          EMBL: AF419301; AAL09704.1;
DR          InterPro: IPR004839; AminoTransf1/2.
DR          InterPro: IPR000796; AsptTransf_sub.
DR          InterPro: IPR004838; Nitransf_1.
DR          Pfam: PF00155; aminotran_1.2; 1.
DR          PRINTS: PR00799; TRANSAMINASE.
DR          PROSITE: PS00105; AA_TRANSFER_CLASS_1; UNKNOWN_1.
KW          Aminotransferase; Transferase.
FT          NON_TER
SQ          SEQUENCE 341 AA; 37396 MW; 915AB17531FD092E CRC64;

Query Match          30.7%; Score 51; DB 10; Length 341;
Best Local Similarity 37.0%; Pred. No. 39;
Matches 10; Conservative 5; Mismatches 10; Indels 2; Gaps 1;

OY          1 SWRSGSFVQGLFGHPDYFYLMEPAW 27
DB          46 SLRVGGEFLGKHY--HQRIVLPQPTW 70

RESULT 40
Q8VP21      PRELIMINARY;      PRT;      681 AA.
ID          Q8VP21;
AC          08VP21;
DT          01-MAR-2002 (TREMBlrel. 20, Created)
DT          01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT          01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE          BcscC.
OS          Burkholderia cepacia genomovar III.
OC          Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC          Burkholderia.
OX          NCBI_TaxID=95486;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          MEDLINE=21441674; PubMed=11557147;
RA          Parsons Y.N., Glendinning K.J., Thornton V., Hales B.A., Hart C.A.,
RT          Winstanley C.;
RT          "A putative type III secretion gene cluster is widely distributed in
RL          the Burkholderia cepacia complex but absent from genomovar I."
DR          FEMS Microbiol. Lett. 203:103-108(2001).
DR          EMBL: AY028431; AAL59397.1;
DR          InterPro: IPR004845; GSP11/IIIProtein.
DR          InterPro: IPR004845; GSP11/IIIProtein.
DR          Pfam: PF00263; GSP11-II; 1.
DR          PROSITE: PS00875; T2SP_D; UNKNOWN_1.
SQ          SEQUENCE 681 AA; 72502 MW; 9C512C0F8A05AE0B CRC64;

Query Match          30.7%; Score 51; DB 2; Length 681;
Best Local Similarity 36.7%; Pred. No. 83;
Matches 11; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

OY          2 WRSGSFVQGLF-----GHPDYFYLMEP 25
DB          568 WLSDIPLIGNLFRYDKSGNHNMERFYLLTP 597

```

```

RESULT 41
OQ25D2 PRELIMINARY; PRT; 158 AA.
AC OQ25D2:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Bcl2-interacting killer-like protein.
GN BIKLK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Itoh T., Itoh A., Pleasure D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF372501; AAK53820.1; -
SQ SEQUENCE 158 AA; 17660 MW; A63742550879C78 CRC64;

Query Match
Best Local Similarity 41.4%; Pred. No. 23;
Matches 12; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

OQ 1 SMRSGSFYGLFGQHPDYFLMEPAHV 29
Db 127 AMVSPDQDBGQLPQVYLVLGLGSAHML 155

RESULT 42
OQVMS5 PRELIMINARY; PRT; 239 AA.
AC OQVMS5:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Hypothetical 26.7 kDa protein.
OS Pseudomonas putida.
OC Pseudomonas.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RA Greated A., Lambertson L., Williams P.A., Thomas C.M.;
RT "Complete nucleotide sequence of IncP-9 plasmid PMW0."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ344068; CAC86851.1;
KM Hypothetical protein; Plasmid.
SQ SEQUENCE 239 AA; 26726 MW; A65732EA71C5B0A6 CRC64;

Query Match
Best Local Similarity 47.4%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OQ 8 FVGLFGQHPDYFLMEPA 26
Db 83 FMGQZGSHEDIDYFVPA 101

RESULT 43
OQJN46 PRELIMINARY; PRT; 678 AA.
AC OQJN46:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative alpha amylase (Fragment).
GN AAM1.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.

```

```

OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=96079285; PubMed=7494484;
RA Ward M.J., Bell A.W., Hamblin P.A., Packer H.L., Armitage J.P.;
RT "Identification of a chemotaxis operon with two cheY genes in
RT Rhodobacter sphaeroides."
RL Mol. Microbiol. 17:357-366(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=96154945; PubMed=8596451;
RA Ward M.J., Harrison D.W., Esher M.J., Armitage J.P.;
RT "Identification of a methyl-accepting chemotaxis protein in
RT Rhodobacter sphaeroides."
RL Mol. Microbiol. 18:115-121(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=20428429; PubMed=10970853;
RA Shah D.S.H., Porter S.L., Martin A.C., Hamblin P.A., Armitage J.P.;
RT "Fine tuning bacterial chemotaxis: Analysis of Rhodobacter sphaeroides
RT behaviour under aerobic and anaerobic conditions by mutation of the
RT major chemotaxis operons and cheY genes."
RL EMBL: J. 19:4601-4613(2000).
DR EMBL: X80205; CAB87126.1;
DR InterPro: IPR000461; Alpha_Amylase.
DR InterPro: IPR000294; Vtk_dep_GLA.
DR Pfam: PF00128; alpha-amylase; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; UNKNOWN_1.
FT NON_TER 678
SQ SEQUENCE 678 AA; 77092 MW; E88BED0339F1B61D CRC64;

Query Match
Best Local Similarity 50.0%; Pred. No. 1,1e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OQ 9 VGLFGQHPDYFLME 24
Db 475 IGEVQSGHPDYFLAE 490

RESULT 44
OQJXN4 PRELIMINARY; PRT; 1051 AA.
AC OQJXN4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE WD repeat domain protein.
GN F14L2-80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordan N., Bangert S., Wiedelmann R., Voss H., Unsel M., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA ED Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AL353818; CAB88535.1;
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 2.
DR SMART; SM00320; WD40; 5.

```

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE; PS50082; WD_REPEATS_2; 3.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 1051 AA; 115562 MW; 881F9E2A826778C9 CRC64;

Query Match
 Best Local Similarity 30.1%; Score 50; DB 10; Length 1051;
 Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 1 SMRSGSFVGOIFGQHPDV--FYLMPEAWH 28
 Db 823 SMNIGSAPGCEIAGLYDVARKYMARKGN 852

RESULT 45
 ID 099140 PRELIMINARY; PRT; 1645 AA.
 AC 099140;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Beta-H spectrin (Fragment).
 GN KST OR CG12008.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN RP
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEAD;
 RX MEDLINE=91035599; PubMed=2229176;
 RA Dubreuil R.R., Byers T.J., Stewart C.T., Kiehart D.P.;
 RT "A beta-spectrin isoform from Drosophila (beta H) is similar in size
 RT to vertebrate dystrophin.";
 RL J. Cell Biol. 111:1849-1858(1990).
 CC -1- FUNCTION: SPECTRIN IS A MAJOR CONSTITUENT OF THE CYOSKELETON
 CC NETWORK. THE NATIVE SPECTRIN FORMED FROM THE BETA-H SUBUNIT,
 CC RESEMBLES A NON-EYTHROID TYPE
 CC -1- FUNCTION: BETA SPECTRIN ALONE MAY BE IMPORTANT FOR RECEPTOR
 CC CLUSTERING.
 CC -1- SUBUNIT: NATIVE SPECTRIN MOLECULE IS A TETRAMER COMPOSED OF TWO
 CC ANTIPARALLEL HETERODIMERS JOINED HEAD TO HEAD SO THAT EACH END OF
 CC THE NATIVE MOLECULE INCLUDES THE C-TERMINUS OF THE ALPHA SUBUNIT
 CC AND THE N-TERMINUS OF THE BETA-H SUBUNIT.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL X53992; CAA37939.1; -.
 DR PIR; A37792; A37792.
 DR PIR; S15666; S15666.
 DR HSSP; Q01082; 1BKR.
 DR FLYBASE; F8an0004167; Kst.
 DR InterPro; IPR001589; Actbind_actnln.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; Spectrin; 1.
 DR Pfam; PF00435; Spectrin; 13.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00150; SPEC; 10.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50002; SH3; 1.
 KM Actin-binding; Capping protein; Cytoskeleton; Membrane; Repeat;
 KM SH3 domain.
 FT NON_TER 1 1
 FT DOMAIN 1 259 SEGMENT 1.
 FT DOMAIN 260 371 SEGMENT 2.
 FT DOMAIN 372 482 SEGMENT 3.

FT DOMAIN 801 961 SEGMENT 7.
 FT REPEAT 483 588 SEGMENT 4.
 FT REPEAT 589 694 SEGMENT 5.
 FT REPEAT 695 800 SEGMENT 6.
 FT REPEAT 962 1063 SEGMENT 8.
 FT REPEAT 1064 1169 SEGMENT 9.
 FT REPEAT 1170 1274 SEGMENT 10.
 FT REPEAT 1275 1380 SEGMENT 11.
 FT REPEAT 1381 1484 SEGMENT 12.
 FT REPEAT 1485 1590 SEGMENT 13.
 FT REPEAT 1591 1645 SEGMENT 14.
 FT NON_TER 1645 1645
 SQ SEQUENCE 1645 AA; 190480 MW; 22FEDCBFBF83023EF CRC64;

Query Match
 Best Local Similarity 45.0%; Score 50; DB 5; Length 1645;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 SMRSGSFVGOIFGQHPDV 20
 Db 186 SMRSGGFNALHSHRPDLF 205

RESULT 46
 ID 09VZ03 PRELIMINARY; PRT; 4097 AA.
 AC 09VZ03; 044084;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE KST protein (Beta-H spectrin).
 GN KST OR KARST OR CG12008.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN RN
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Adamantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brodstein P., Brothier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jaitali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Krat C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon S., Nuskern D.R., Paley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE OF 1633-2603 FROM N.A.
 RA Thomas G.H., Newbern E.C., Korte C.C., Bales M.A., Muse S.V.,
 RA Clark A.G., Klehart D.P.;
 RL Mol. Biol. Evol. 0:0-0(1997).
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: AE003477; AAR47766.1; -;
 DR EMBL: AF022656; AAB87867.1; -;
 DR HSSP: Q01082; 1BKR.
 DR FLYBASE: FBgn0004167; kst.
 DR InterPro: IPR001589; Actbind_actin.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR001605; Spectrin_PH.
 DR Pfam: PF00307; CH; 2.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00435; Spectrin; 31.
 DR PRINTS: PR00663; SPECTRINPH.
 DR PRODOM: PD000066; SH3; 1.
 DR SMART: SM00033; CH; 2.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00150; SPEC; 29.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00021; CH; 2.
 DR PROSITE: PS50002; PH_DOMAIN; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain.
 SQ SEQUENCE 4097 AA; 471636 MW; EDE01C3857FCA636 CRC64;
 Query Match 30.1%; Score 50; DB 5; Length 4097;
 Best Local Similarity 45.0%; Pred. No. 8.2e+02;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 QY 1 SWRSSSFVQGLFGQHPDVF 20
 DB 173 SWRSSGLGFNALIHSHRPLF 192
 ||||| : ||:|
 RESULT 47
 Q91540 PRELIMINARY; PRT; 634 AA.
 AC Q91540;
 DT 01-MAR-2001 (TReMBLrel. 15, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Potassium uptake protein KUP.
 GN KUP OR PA0917.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NCBI_TaxID=287;
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gaeher R.L., Goltzy L., Tolentino E., Westbrock-Madman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reller J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an

RT Opportunistic pathogen,"
 RL Nature 406:959-964(2000).
 DR EMBL: AE004526; AAG04306.1; -;
 DR InterPro: IPR003855; K+transprtr.
 DR Pfam: PF02705; Ktrans; 1.
 KW Complete proteome.
 SQ SEQUENCE 634 AA; 68734 MW; 5D4CBA5F70CCBE99 CRC64;
 Query Match 29.8%; Score 49.5; DB 16; Length 634;
 Best Local Similarity 26.1%; Pred. No. 1.2e+02;
 Matches 12; Conservative 5; Mismatches 10; Indels 19; Gaps 1;
 QY 3 RSSGSFVQGLFG-----QHPDVFYIMEPA 29
 DB 171 KHGTARIGILFGPYWIMFGALALGVYGIQPEVYLQANPMWAV 216
 ||:|:| ||||| : |||||
 RESULT 48
 Q9NV23 PRELIMINARY; PRT; 265 AA.
 AC Q9NV23;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE CDNA FLJ10982 fis, clone PLACE1001692, moderately similar to S-acyl
 DE fatty acid synthase thioesterase, medium chain (Ec 3.1.2.14).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hata H.,
 RA Tanase T., Nomura Y., Togiya S., Komai Y., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Oshima A.;
 RT "MDO human CDNA sequencing project,"
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK001844; BA091937.1; -;
 DR InterPro: IPR000379; Ser_estrs_site.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00975; Thioesterase; 1.
 SQ SEQUENCE 265 AA; 29931 MW; C67722F3CAD2D2C7 CRC64;
 Query Match 29.5%; Score 49; DB 4; Length 265;
 Best Local Similarity 41.4%; Pred. No. 56;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 QY 1 SWRSSSFVQGLFGQHPDVFYIMEPA 26
 DB 219 AMKDVTSGNNAKIYQLPFGH---FYLLDPA 244
 ||:|:| ||||| : |||||
 RESULT 49
 Q9NW1 PRELIMINARY; PRT; 318 AA.
 AC Q9NW1;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE CDNA FLJ11106 fis, clone PLACE1005763, moderately similar to S-acyl
 DE fatty acid synthase thioesterase, medium chain (Ec 3.1.2.14).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

RA Matsunawa H., Ishii S., Kawai Y., Saigo K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK001968; BAA92007.1; -;
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00975; Thioesterase; 1.
 SQ SEQUENCE 318 AA; 35818 MW; C7DB0DFEBCB0823F CRC64;

Query Match 29.5%; Score 49; DB 4; Length 318;
 Best Local Similarity 41.4%; Pred. No. 69;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSSFFVGLFGQHPDVFLMEPA 26
 Db 272 AKKDVITSGNAKIYQLPFGH--FYLDPFA 297

RESULT 50

095UT2 PRELIMINARY; PRT; 891 AA.

AC 095UT2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Iron regulatory protein 1.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Spingioidea; Spingidae; Spinginae; Manduca.
 OX NCBI_TaxID=7130;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang D., Ferris C., Gailer J., Kohlhepp P., Winzerling J.J.;
 RT "Manduca sexta IRPI: molecular characterization and in vivo response
 to iron";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032658; AAK39637.1; -;
 DR InterPro: IPR000573; Aconitase_C.
 DR InterPro: IPR001030; Aconitase_N.
 DR Pfam: PF00330; aconitase; 1.
 DR Pfam: PF00694; aconitase; 1.
 DR ProDom: PD000511; aconitase_N; 1.
 DR PROSITE; PS00450; ACONITASE_1; UNKNOWN_1.
 DR PROSITE; PS01244; ACONITASE_2; UNKNOWN_1.
 SQ SEQUENCE 891 AA; 97268 MW; 592D36D150E041E4 CRC64;

Query Match 29.5%; Score 49; DB 5; Length 891;
 Best Local Similarity 40.0%; Pred. No. 21e+02;
 Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFLMEP 25
 Db 626 SMOGLSVPGKLYGWDPNSTYIKKP 650

Search completed: February 20, 2003, 13:33:24
 Job time : 73.4286 secs

This page Blank (uspio)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:31 ; Search time 33.1429 Seconds
(without alignments)
25.745 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166

Sequence: 1 SWRSGSFVGOLEFGHPDYFIMEPMNV 29

Scoring table: BLOSUM62
Gapol 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Issued Patents AA:*

1: /cgn2_6/p/locdata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/p/locdata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/p/locdata/1/1aa/6A.COMB.pep:*

4: /cgn2_6/p/locdata/1/1aa/6B.COMB.pep:*

5: /cgn2_6/p/locdata/1/1aa/PCRTUS.COMB.pep:*

6: /cgn2_6/p/locdata/1/1aa/Backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	166	100.0	US-09-045-284A-2	Sequence 2, Appli
2	166	100.0	US-09-190-911-1	Sequence 1, Appli
3	124	74.7	US-09-263-023-2	Sequence 2, Appli
4	124	74.7	US-09-471-867-2	Sequence 2, Appli
5	124	74.7	US-09-263-023-4	Sequence 4, Appli
6	124	74.7	US-09-471-867-4	Sequence 4, Appli
7	116	69.9	US-09-015-188-2	Sequence 2, Appli
8	101	60.8	US-08-655-878-2	Sequence 2, Appli
9	101	60.8	US-08-899-514-2	Sequence 2, Appli
10	53.5	32.2	US-08-443-865-6	Sequence 6, Appli
11	53.5	32.2	US-08-900-491-6	Sequence 6, Appli
12	47	28.3	US-09-202-491-7	Sequence 7, Appli
13	46.5	28.0	US-08-737-825-6	Sequence 6, Appli
14	45	27.1	US-09-134-001C-5331	Sequence 5331, Ap
15	45	27.1	US-08-551-386-4	Sequence 4, Appli
16	45	27.1	PCT-US93-12687-4	Sequence 4, Appli
17	45	27.1	US-08-206-176-2	Sequence 2, Appli
18	45	27.1	US-09-643-597-170	Sequence 170, App
19	45	27.1	US-09-373-157-4	Sequence 4, Appli
20	45	27.1	US-09-643-597-357	Sequence 357, App
21	45	27.1	US-09-193-562D-32	Sequence 32, Appli
22	45	27.1	US-09-643-597-161	Sequence 161, App
23	45	27.1	US-08-551-386-6	Sequence 6, Appli
24	45	27.1	PCT-US93-12687-2	Sequence 6, Appli
25	44	26.5	US-09-357-251-22	Sequence 22, Appli
26	44	26.5	US-09-052-877-2	Sequence 2, Appli
27	44	26.5	US-08-740-223A-15	Sequence 15, Appli

28	44	26.5	496	US-09-709-188-15	Sequence 15, Appli
29	44	26.5	627	US-08-295-814E-4	Sequence 4, Appli
30	44	26.5	627	US-08-291-299-10	Sequence 10, Appli
31	44	26.5	627	US-08-343-361-4	Sequence 4, Appli
32	44	26.5	627	PCT-US93-01959-14	Sequence 14, Appli
33	44	26.5	627	PCT-US93-01959-10	Sequence 10, Appli
34	44	26.5	632	US-08-295-814E-10	Sequence 10, Appli
35	44	26.5	632	US-09-343-361-10	Sequence 10, Appli
36	44	26.5	632	PCT-US93-01959-10	Sequence 10, Appli
37	44	26.5	1271	US-08-095-734-2	Sequence 2, Appli
38	44	26.5	1271	US-08-444-623-2	Sequence 2, Appli
39	44	26.5	1271	US-08-471-869-2	Sequence 2, Appli
40	44	26.5	1271	US-09-342-563-2	Sequence 2, Appli
41	44	26.5	1271	PCT-US94-08267-2	Sequence 2, Appli
42	44	26.5	2509	US-08-469-005A-10	Sequence 10, Appli
43	44	26.5	2511	US-09-261-907-2	Sequence 2, Appli
44	43	25.9	256	US-08-906-769-89	Sequence 89, Appli
45	43	25.9	256	US-08-906-616-89	Sequence 89, Appli
46	43	25.9	256	US-08-817-795-89	Sequence 89, Appli
47	43	25.9	256	US-08-639-075A-89	Sequence 89, Appli
48	43	25.9	256	US-09-012-431-89	Sequence 89, Appli
49	43	25.9	256	US-09-012-692-89	Sequence 89, Appli
50	43	25.9	256	US-08-906-613-89	Sequence 89, Appli
51	43	25.9	256	PCT-US95-1442A-88	Sequence 89, Appli
52	43	25.9	467	US-09-134-001C-4200	Sequence 4200, Ap
53	42	25.3	22	US-08-003-966-3	Sequence 3, Appli
54	42	25.3	235	US-08-776-059-39	Sequence 39, Appli
55	42	25.3	250	US-08-872-651A-1	Sequence 1, Appli
56	42	25.3	250	US-09-231-258-1	Sequence 1, Appli
57	42	25.3	253	US-08-776-059-31	Sequence 31, Appli
58	42	25.3	503	US-09-357-251-36	Sequence 36, Appli
59	42	25.3	504	US-08-776-059-35	Sequence 35, Appli
60	42	25.3	656	US-09-605-785-379	Sequence 379, App
61	42	25.3	656	US-09-439-313-379	Sequence 379, App
62	42	25.3	656	US-09-352-616A-379	Sequence 379, App
63	42	25.3	671	US-09-605-785-380	Sequence 380, App
64	42	25.3	671	US-09-439-313-380	Sequence 380, App
65	42	25.3	671	US-09-352-616A-380	Sequence 380, App
66	42	25.3	1719	US-09-605-785-378	Sequence 378, App
67	42	25.3	1719	US-09-439-313-378	Sequence 378, App
68	42	25.3	1719	US-09-352-616A-378	Sequence 378, App
69	41.5	25.0	702	US-09-134-001C-2842	Sequence 2842, Ap
70	41.5	25.0	702	US-09-232-200-102	Sequence 102, App
71	41.5	25.0	702	US-09-232-197-102	Sequence 102, App
72	41.5	25.0	702	US-09-232-201-102	Sequence 102, App
73	41	24.7	80	US-08-485-455D-55	Sequence 55, Appli
74	41	24.7	80	US-08-482-130C-55	Sequence 55, Appli
75	41	24.7	80	US-08-484-211C-55	Sequence 55, Appli
76	41	24.7	80	US-08-906-769-55	Sequence 55, Appli
77	41	24.7	80	US-08-906-616-55	Sequence 55, Appli
78	41	24.7	80	US-08-817-795-55	Sequence 55, Appli
79	41	24.7	80	US-08-485-443B-55	Sequence 55, Appli
80	41	24.7	80	US-08-639-075A-55	Sequence 55, Appli
81	41	24.7	80	US-09-012-431-55	Sequence 55, Appli
82	41	24.7	80	US-09-012-692-55	Sequence 55, Appli
83	41	24.7	80	US-08-906-613-55	Sequence 55, Appli
84	41	24.7	80	PCT-US95-1442A-55	Sequence 55, Appli
85	41	24.7	189	US-07-982-650D-2	Sequence 2, Appli
86	41	24.7	189	US-08-331-379-3	Sequence 3, Appli
87	41	24.7	189	US-07-982-650D-1	Sequence 1, Appli
88	41	24.7	189	US-07-982-650D-3	Sequence 3, Appli
89	41	24.7	189	US-07-982-650D-4	Sequence 4, Appli
90	41	24.7	189	US-07-982-650D-5	Sequence 5, Appli
91	41	24.7	189	US-08-331-379-1	Sequence 1, Appli
92	41	24.7	189	US-08-331-379-3	Sequence 3, Appli
93	41	24.7	189	US-08-331-379-4	Sequence 4, Appli
94	41	24.7	189	US-08-331-379-5	Sequence 5, Appli
95	41	24.7	226	US-09-032-515-32	Sequence 32, Appli
96	41	24.7	226	US-08-549-515-7	Sequence 7, Appli
97	41	24.7	349	US-08-465-971B-2	Sequence 2, Appli
98	41	24.7	349	US-08-806-597A-14	Sequence 14, Appli
99	41	24.7	349	US-08-970-028A-14	Sequence 14, Appli
100	41	24.7	393	US-08-961-083-58	Sequence 58, Appli

101	41	24.7	429	4	US-09-194-905-9	Sequence 9, Appl	174	39	23.5	17	4	US-09-990-833-91	Sequence 91, Appl
102	41	24.7	599	4	US-09-000-145-2	Sequence 2, Appl	175	39	23.5	32	4	US-09-149-476-442	Sequence 442, App
103	41	24.7	686	4	US-09-653-839-8	Sequence 8, Appl	176	39	23.5	133	4	US-08-905-223-112	Sequence 412, App
104	41	24.7	702	4	US-09-653-839-6	Sequence 6, Appl	177	39	23.5	198	4	US-09-045-973-5	Sequence 5, Appl
105	41	24.7	723	4	US-09-653-839-4	Sequence 4, Appl	178	39	23.5	231	4	US-09-134-001C-4404	Sequence 4404, Ap
106	41	24.7	739	4	US-09-653-839-2	Sequence 2, Appl	179	39	23.5	244	3	US-08-975-658A-24	Sequence 24, Appl
107	40.5	24.4	497	4	US-09-594-193-2	Sequence 2, Appl	180	39	23.5	244	3	US-08-975-658A-28	Sequence 28, Appl
108	40.5	24.4	5087	4	US-09-144-085-1	Sequence 1, Appl	181	39	23.5	244	4	US-09-417-090-28	Sequence 28, Appl
109	40.5	24.4	5215	4	US-09-105-537-2	Sequence 2, Appl	182	39	23.5	244	4	US-09-727-578-28	Sequence 28, Appl
110	40.5	24.4	5588	4	US-09-036-987A-6	Sequence 6, Appl	183	39	23.5	282	1	US-07-712-476A-5	Sequence 5, Appl
111	40.5	24.4	5588	4	US-09-370-700-6	Sequence 6, Appl	184	39	23.5	294	4	US-09-424-349A-3	Sequence 3, Appl
112	40	24.1	17	4	US-08-990-823-90	Sequence 90, Appl	185	39	23.5	295	4	US-09-134-001C-5643	Sequence 5643, Ap
113	40	24.1	126	4	US-09-288-143-71	Sequence 71, Appl	186	39	23.5	296	1	US-07-712-476A-1	Sequence 1, Appl
114	40	24.1	238	4	US-09-134-001C-3136	Sequence 3136, Ap	187	39	23.5	300	6	5340934-6	Patent No. 5340934
115	40	24.1	315	1	US-08-129-129-8	Sequence 8, Appl	188	39	23.5	318	4	US-09-355-166-16	Sequence 16, Appl
116	40	24.1	347	1	US-08-445-515-58	Sequence 58, Appl	189	39	23.5	370	3	US-09-150-133-1	Sequence 1, Appl
117	40	24.1	348	3	US-08-445-515-56	Sequence 56, Appl	190	39	23.5	370	3	US-09-150-133-3	Sequence 3, Appl
118	40	24.1	359	3	US-09-150-133-11	Sequence 11, Appl	191	39	23.5	370	3	US-09-150-141-1	Sequence 1, Appl
119	40	24.1	359	3	US-09-150-141-11	Sequence 11, Appl	192	39	23.5	370	3	US-09-150-141-3	Sequence 3, Appl
120	40	24.1	359	4	US-09-374-493-11	Sequence 11, Appl	193	39	23.5	370	4	US-09-374-493-1	Sequence 1, Appl
121	40	24.1	359	4	US-09-374-824-11	Sequence 11, Appl	194	39	23.5	370	4	US-09-374-493-3	Sequence 3, Appl
122	40	24.1	385	4	US-08-891-492-11	Sequence 11, Appl	195	39	23.5	370	4	US-09-374-824-1	Sequence 1, Appl
123	40	24.1	385	1	US-08-891-254-3	Sequence 3, Appl	196	39	23.5	370	4	US-09-374-824-3	Sequence 3, Appl
124	40	24.1	385	5	US-08-819-539-3	Sequence 3, Appl	197	39	23.5	370	4	US-09-374-492-1	Sequence 1, Appl
125	40	24.1	385	5	PCT-US93-06243-2	Sequence 2, Appl	198	39	23.5	370	4	US-09-374-492-3	Sequence 3, Appl
126	40	24.1	385	5	PCT-US96-08819-3	Sequence 3, Appl	199	39	23.5	513	4	US-07-936-163-4	Sequence 4, Appl
127	40	24.1	403	2	US-08-200-724A-2	Sequence 2, Appl	200	39	23.5	583	4	US-09-739-455-2	Sequence 2, Appl
128	40	24.1	403	2	US-09-030-270A-3	Sequence 3, Appl	201	39	23.5	586	3	US-09-739-455-2	Sequence 3, Appl
129	40	24.1	403	4	US-08-851-376A-2	Sequence 2, Appl	202	39	23.5	587	3	US-08-964-268-3	Sequence 3, Appl
130	40	24.1	403	4	US-08-984-207-3	Sequence 3, Appl	203	39	23.5	711	4	US-08-961-083-82	Sequence 82, Appl
131	40	24.1	403	4	US-09-013-587-3	Sequence 3, Appl	204	39	23.5	737	3	US-09-163-356-2	Sequence 2, Appl
132	40	24.1	461	2	US-08-527-227A-7	Sequence 7, Appl	205	39	23.5	1337	3	US-08-854-585-2	Sequence 2, Appl
133	40	24.1	592	6	5200183-18	Patent No. 5200183	206	39	23.5	1632	5	PCT-US95-05512-2	Sequence 2, Appl
134	40	24.1	685	3	US-08-947-965-74	Sequence 74, Appl	207	39	23.5	1684	3	US-08-665-259-35	Sequence 25, Appl
135	40	24.1	788	2	US-08-907-166-6	Sequence 6, Appl	208	39	23.5	1684	3	US-08-762-500-55	Sequence 25, Appl
136	40	24.1	940	1	US-08-938-365-4	Sequence 4, Appl	209	39	23.5	1704	3	US-08-762-500-75	Sequence 75, Appl
137	40	24.1	941	1	US-08-943-760A-2	Sequence 2, Appl	210	39	23.5	233	4	US-08-199-637A-81	Sequence 81, Appl
138	40	24.1	2154	2	US-08-841-349-4	Sequence 4, Appl	211	38.5	23.2	436	2	US-08-576-626A-47	Sequence 47, Appl
139	39.5	23.8	52	1	US-08-485-4550-79	Sequence 79, Appl	212	38.5	23.2	462	4	US-09-036-987A-18	Sequence 18, Appl
140	39.5	23.8	52	2	US-08-482-130C-79	Sequence 79, Appl	213	38.5	23.2	462	4	US-09-370-700-18	Sequence 18, Appl
141	39.5	23.8	52	2	US-08-484-211C-79	Sequence 79, Appl	214	38.5	23.2	472	2	US-08-622-166A-2	Sequence 2, Appl
142	39.5	23.8	52	3	US-08-906-769-79	Sequence 79, Appl	215	38.5	23.2	472	2	US-08-622-166A-4	Sequence 4, Appl
143	39.5	23.8	52	3	US-08-906-616-79	Sequence 79, Appl	216	38.5	23.2	543	1	US-08-362-232-2	Sequence 2, Appl
144	39.5	23.8	52	4	US-08-817-795-79	Sequence 79, Appl	217	38.5	23.2	543	1	US-08-814-186-2	Sequence 2, Appl
145	39.5	23.8	52	4	US-08-485-443B-79	Sequence 79, Appl	218	38.5	23.2	566	2	US-07-862-588B-4	Sequence 4, Appl
146	39.5	23.8	52	4	US-08-639-075A-79	Sequence 79, Appl	219	38.5	23.2	1150	4	US-09-238-303-9	Sequence 9, Appl
147	39.5	23.8	52	4	US-09-012-431-79	Sequence 79, Appl	220	38.5	23.2	3224	2	US-08-705-660-34	Sequence 34, Appl
148	39.5	23.8	52	4	US-09-012-692-79	Sequence 79, Appl	221	38.5	23.2	3224	3	US-08-989-045-14	Sequence 34, Appl
149	39.5	23.8	52	4	US-08-906-613-79	Sequence 79, Appl	222	38.5	23.2	3665	2	US-08-222-617A-13	Sequence 13, Appl
150	39.5	23.8	52	5	PCT-US95-14442A-79	Sequence 79, Appl	223	38.5	23.2	3712	2	US-08-222-617A-4	Sequence 4, Appl
151	39.5	23.8	252	3	US-08-906-769-103	Sequence 103, App	224	38.5	23.2	3712	2	US-08-222-617A-25	Sequence 25, Appl
152	39.5	23.8	252	4	US-08-906-616-103	Sequence 103, App	225	38.5	23.2	466	1	US-07-826-945A-8	Sequence 8, Appl
153	39.5	23.8	252	4	US-08-817-795-103	Sequence 103, App	226	38.5	23.2	169	4	US-09-188-930-150	Sequence 150, App
154	39.5	23.8	252	4	US-08-639-075A-103	Sequence 103, App	227	38	22.9	169	4	US-09-188-930-164	Sequence 164, App
155	39.5	23.8	252	4	US-09-012-431-103	Sequence 103, App	228	38	22.9	169	4	US-09-188-930-282	Sequence 282, App
156	39.5	23.8	252	4	US-09-012-692-103	Sequence 103, App	229	38	22.9	169	4	US-09-188-930-287	Sequence 287, App
157	39.5	23.8	252	4	US-08-906-613-103	Sequence 103, App	230	38	22.9	216	4	US-09-077-955-31	Sequence 31, Appl
158	39.5	23.8	252	5	PCT-US95-14442A-103	Sequence 103, App	231	38	22.9	229	4	US-09-077-955-30	Sequence 30, Appl
159	39.5	23.8	333	2	US-08-044-812A-4	Sequence 4, Appl	232	38	22.9	256	4	US-09-077-955-7	Sequence 7, Appl
160	39.5	23.8	333	3	US-08-475-637-4	Sequence 4, Appl	233	38	22.9	273	2	US-07-857-224B-73	Sequence 73, Appl
161	39.5	23.8	333	3	US-08-706-281A-12	Sequence 12, Appl	234	38	22.9	277	2	US-08-469-537A-57	Sequence 57, Appl
162	39.5	23.8	333	4	US-09-191-359-4	Sequence 4, Appl	235	38	22.9	284	4	US-09-357-251-24	Sequence 24, Appl
163	39.5	23.8	337	4	US-09-097-231-12	Sequence 12, Appl	236	38	22.9	294	4	US-09-077-955-29	Sequence 29, Appl
164	39.5	23.8	337	3	US-09-320-878-15	Sequence 15, Appl	237	38	22.9	330	2	US-08-838-553-3	Sequence 3, Appl
165	39.5	23.8	337	3	US-09-105-537-14	Sequence 14, Appl	238	38	22.9	338	2	US-08-891-254-1	Sequence 1, Appl
166	39.5	23.8	487	4	US-09-594-193-13	Sequence 13, Appl	239	38	22.9	338	2	US-08-484-358-2	Sequence 2, Appl
167	39.5	23.8	632	4	US-09-354-129-8	Sequence 8, Appl	240	38	22.9	338	2	US-08-819-539-1	Sequence 1, Appl
168	39.5	23.8	632	4	US-09-504-357-8	Sequence 8, Appl	241	38	22.9	338	2	US-09-030-270A-1	Sequence 1, Appl
169	39.5	23.8	688	1	US-08-205-018-2	Sequence 2, Appl	242	38	22.9	338	2	US-09-118-959-2	Sequence 2, Appl
170	39.5	23.8	859	1	US-08-995-580-2	Sequence 2, Appl	243	38	22.9	338	4	US-08-984-207-1	Sequence 1, Appl
171	39.5	23.8	877	5	PCT-US95-02792-4	Sequence 4, Appl	244	38	22.9	338	4	US-09-013-587-1	Sequence 1, Appl
172	39.5	23.8	877	1	US-08-430-925A-4	Sequence 4, Appl	245	38	22.9	338	4	US-09-077-955-28	Sequence 28, Appl
173	39.5	23.8	3782	4	US-09-105-537-4	Sequence 4, Appl	246	38	22.9	338	4		

247	38	22.9	338	5	PCT-US96-08819-1	Sequence 1, Appl.	320	37	22.3	753	4	US-08-026-143B-3	Sequence 3, Appl.
248	38	22.9	390	4	US-09-077-955-27	Sequence 27, Appl	321	37	22.3	753	5	PCT-US99-10621-3	Sequence 3, Appl.
249	38	22.9	440	4	US-09-077-955-26	Sequence 26, Appl	322	37	22.3	753	5	PCT-US94-00233-3	Sequence 3, Appl.
250	38	22.9	456	4	US-09-077-955-25	Sequence 25, Appl	323	37	22.3	788	1	US-08-346-455B-36	Sequence 36, Appl
251	38	22.9	492	2	US-08-644-271-32	Sequence 32, Appl	324	37	22.3	788	3	US-08-977-221-36	Sequence 36, Appl
252	38	22.9	492	4	US-09-077-955-36	Sequence 36, Appl	325	37	22.3	788	4	US-09-483-831B-36	Sequence 36, Appl
253	38	22.9	499	4	US-08-988-856B-2	Sequence 2, Appl	326	37	22.3	788	5	PCT-US95-06613-36	Sequence 36, Appl
254	38	22.9	499	5	PCT-US95-01806-2	Sequence 2, Appl	327	37	22.3	802	4	US-09-173-151A-33	Sequence 33, Appl
255	38	22.9	520	4	US-08-961-083-122	Sequence 122, App	328	37	22.3	829	1	US-08-346-455B-34	Sequence 34, App
256	38	22.9	522	4	US-09-134-001C-3755	Sequence 3755, Ap	329	37	22.3	829	3	US-08-977-221-34	Sequence 34, App
257	38	22.9	545	4	US-08-936-107A-10	Sequence 10, Appl	330	37	22.3	829	4	US-09-483-831B-34	Sequence 34, App
258	38	22.9	553	4	US-08-997-251-2	Sequence 2, Appl	331	37	22.3	829	5	PCT-US99-06613-34	Sequence 34, App
259	38	22.9	666	3	US-08-905-865-4	Sequence 4, Appl	332	37	22.3	861	1	US-08-346-455B-67	Sequence 67, App
260	38	22.9	696	4	US-09-129-666-4	Sequence 4, Appl	333	37	22.3	861	3	US-08-977-221-67	Sequence 67, App
261	38	22.9	699	4	US-09-134-001C-4054	Sequence 4054, Ap	334	37	22.3	861	4	US-09-483-831B-67	Sequence 67, App
262	38	22.9	973	3	US-08-904-452-4	Sequence 4, Appl	335	37	22.3	861	5	PCT-US95-06613-67	Sequence 67, App
263	38	22.9	973	4	US-09-517-633-4	Sequence 4, Appl	336	37	22.3	915	1	US-08-346-455B-69	Sequence 69, App
264	38	22.9	1040	4	US-08-961-083-118	Sequence 118, App	337	37	22.3	915	3	US-08-977-221-69	Sequence 69, App
265	38	22.9	1248	4	US-08-882-046-6	Sequence 6, Appl	338	37	22.3	915	5	PCT-US95-06613-69	Sequence 69, App
266	38	22.9	1253	1	US-08-252-966B-12	Sequence 12, Appl	339	37	22.3	915	5	PCT-US99-06613-69	Sequence 69, App
267	38	22.9	1261	1	US-08-252-966B-18	Sequence 18, Appl	340	37	22.3	979	1	US-08-346-455B-38	Sequence 38, App
268	38	22.9	1265	4	US-08-469-260A-401	Sequence 401, App	341	37	22.3	979	3	US-08-977-221-38	Sequence 38, App
269	37.5	22.6	391	4	US-09-428-589-2	Sequence 2, Appl	342	37	22.3	979	5	PCT-US95-06613-38	Sequence 38, App
270	37.5	22.6	488	2	US-08-797-226-2	Sequence 2, Appl	343	37	22.3	979	5	PCT-US99-06613-38	Sequence 38, App
271	37.5	22.6	590	2	US-08-756-317-12	Sequence 12, Appl	344	37	22.3	1308	2	US-08-966-644-2	Sequence 2, App
272	37.5	22.6	1174	2	US-08-446-345-36	Sequence 36, Appl	345	37	22.3	1308	3	US-08-352-552-2	Sequence 2, App
273	37	22.3	129	2	US-08-785-065-10	Sequence 10, Appl	346	37	22.3	1430	3	US-09-008-172-2	Sequence 2, App
274	37	22.3	129	4	US-09-151-412-10	Sequence 10, Appl	347	37	22.3	1430	4	US-09-210-361-6	Sequence 6, App
275	37	22.3	209	6	547917-2	Patent No. 547917	348	37	22.3	1430	4	US-09-740-374-6	

393	36.5	22.0	377	4	US-09-374-824-5	Sequence 5, Appl1	466	36.5	22.0	832	2	US-08-599-491-4	Sequence 4, Appl1
394	36.5	22.0	377	4	US-09-374-492-5	Sequence 5, Appl1	467	36.5	22.0	832	2	US-08-756-386-4	Sequence 4, Appl1
395	36.5	22.0	525	3	US-08-888-949-16	Sequence 16, Appl	468	36.5	22.0	832	2	US-08-823-516-4	Sequence 4, Appl1
396	36.5	22.0	525	4	US-08-888-950-16	Sequence 16, Appl	469	36.5	22.0	832	3	US-08-682-853A-4	Sequence 4, Appl1
397	36.5	22.0	525	4	US-09-262-758-16	Sequence 16, Appl	470	36.5	22.0	832	3	US-08-759-038-4	Sequence 4, Appl1
398	36.5	22.0	527	1	US-08-805-814-10	Sequence 10, Appl	471	36.5	22.0	832	3	US-08-759-038-4	Sequence 4, Appl1
399	36.5	22.0	527	1	US-08-484-815-10	Sequence 10, Appl	472	36.5	22.0	832	4	US-09-350-309-4	Sequence 4, Appl1
400	36.5	22.0	527	3	US-08-888-949-10	Sequence 10, Appl	473	36.5	22.0	832	4	US-08-520-946-4	Sequence 4, Appl1
401	36.5	22.0	527	3	US-08-888-950-10	Sequence 10, Appl	474	36.5	22.0	832	4	US-08-978-806-2	Sequence 2, Appl1
402	36.5	22.0	527	4	US-09-262-758-10	Sequence 10, Appl	475	36.5	22.0	832	5	PCT-US91-07035-2	Sequence 2, Appl1
403	36.5	22.0	527	5	PCT-US95-10284-10	Sequence 10, Appl	476	36.5	22.0	833	1	US-08-073-384C-8	Sequence 8, Appl1
404	36.5	22.0	528	2	US-08-484-956-90	Sequence 90, Appl	477	36.5	22.0	833	1	US-08-254-359A-8	Sequence 8, Appl1
405	36.5	22.0	528	2	US-08-757-653-90	Sequence 90, Appl	478	36.5	22.0	833	1	US-08-483-043-8	Sequence 8, Appl1
406	36.5	22.0	528	2	US-08-520-946-90	Sequence 90, Appl	479	36.5	22.0	833	1	US-08-482-228-8	Sequence 8, Appl1
407	36.5	22.0	545	4	US-09-269-731-8	Sequence 86, Appl1	480	36.5	22.0	833	2	US-08-471-066B-8	Sequence 8, Appl1
408	36.5	22.0	548	2	US-08-484-956-86	Sequence 86, Appl1	481	36.5	22.0	833	2	US-08-484-956-8	Sequence 8, Appl1
409	36.5	22.0	548	2	US-08-757-653-86	Sequence 86, Appl1	482	36.5	22.0	833	2	US-08-484-956-85	Sequence 8, Appl1
410	36.5	22.0	548	4	US-08-520-946-86	Sequence 86, Appl1	483	36.5	22.0	833	2	US-08-757-653-8	Sequence 8, Appl1
411	36.5	22.0	582	4	US-09-091-725-17	Sequence 17, Appl	484	36.5	22.0	833	2	US-08-757-653-85	Sequence 8, Appl1
412	36.5	22.0	591	4	US-09-134-001C-3857	Sequence 3857, Ap	485	36.5	22.0	833	2	US-08-757-653-168	Sequence 168, App
413	36.5	22.0	609	4	US-09-232-200-69	Sequence 69, Appl	486	36.5	22.0	833	2	US-08-757-653-188	Sequence 188, App
414	36.5	22.0	609	4	US-09-232-197-69	Sequence 69, Appl	487	36.5	22.0	833	2	US-08-757-653-190	Sequence 190, App
415	36.5	22.0	609	4	US-09-232-201-69	Sequence 69, Appl	488	36.5	22.0	833	2	US-08-599-491-8	Sequence 8, Appl1
416	36.5	22.0	613	4	US-09-232-191-7	Sequence 7, Appl1	489	36.5	22.0	833	2	US-08-756-386-8	Sequence 8, Appl1
417	36.5	22.0	613	4	US-09-232-200-7	Sequence 7, Appl1	490	36.5	22.0	833	2	US-08-823-516-8	Sequence 8, Appl1
418	36.5	22.0	613	4	US-09-232-200-94	Sequence 94, Appl1	491	36.5	22.0	833	2	US-08-823-516-66	Sequence 66, Appl1
419	36.5	22.0	613	4	US-09-232-197-7	Sequence 7, Appl1	492	36.5	22.0	833	2	US-08-823-516-69	Sequence 69, Appl1
420	36.5	22.0	613	4	US-09-232-197-94	Sequence 94, Appl1	493	36.5	22.0	833	2	US-08-823-516-71	Sequence 71, Appl1
421	36.5	22.0	613	4	US-09-232-201-7	Sequence 7, Appl1	494	36.5	22.0	833	3	US-08-682-853A-8	Sequence 8, Appl1
422	36.5	22.0	613	4	US-09-232-201-94	Sequence 94, Appl1	495	36.5	22.0	833	3	US-08-759-038-8	Sequence 8, Appl1
423	36.5	22.0	638	2	US-08-557-122A-38	Sequence 38, Appl	496	36.5	22.0	833	3	US-08-759-038-107	Sequence 107, App
424	36.5	22.0	638	2	US-09-262-666-38	Sequence 38, Appl	497	36.5	22.0	833	3	US-08-759-038-132	Sequence 132, App
425	36.5	22.0	655	4	US-09-232-191-31	Sequence 31, Appl	498	36.5	22.0	833	3	US-08-756-314-8	Sequence 8, Appl1
426	36.5	22.0	655	4	US-09-232-200-81	Sequence 81, Appl	499	36.5	22.0	833	3	US-08-756-314-107	Sequence 107, App
427	36.5	22.0	655	4	US-09-232-197-81	Sequence 81, Appl	500	36.5	22.0	833	3	US-08-756-314-130	Sequence 130, App
428	36.5	22.0	655	4	US-09-232-201-81	Sequence 81, Appl	501	36.5	22.0	833	3	US-08-756-314-132	Sequence 132, App
429	36.5	22.0	664	6	5240838-5	Patent No. 5240838	502	36.5	22.0	833	4	US-08-520-946-8	Sequence 8, Appl1
430	36.5	22.0	695	2	US-08-484-956-87	Sequence 87, Appl	503	36.5	22.0	833	4	US-08-520-946-85	Sequence 85, Appl1
431	36.5	22.0	695	2	US-08-757-653-87	Sequence 87, Appl	504	36.5	22.0	833	4	US-08-352-159-25	Sequence 25, Appl
432	36.5	22.0	695	4	US-08-520-946-87	Sequence 87, Appl	505	36.5	22.0	833	4	US-09-352-168-35	Sequence 35, Appl
433	36.5	22.0	715	4	US-09-620-412C-329	Sequence 329, App	506	36.5	22.0	1000	4	US-09-352-168-35	Sequence 35, Appl
434	36.5	22.0	810	4	US-09-587-856-2	Sequence 2, Appl1	507	36.5	22.0	1000	4	US-08-466-537A-96	Sequence 96, Appl
435	36.5	22.0	810	4	US-09-777-537-2	Sequence 2, Appl1	508	36.5	22.0	1101	2	US-07-934-338B-2	Sequence 2, Appl1
436	36.5	22.0	810	4	US-09-777-538-2	Sequence 2, Appl1	509	36.5	22.0	1118	1	US-08-278-089A-2	Sequence 2, Appl1
437	36.5	22.0	831	1	US-08-073-384C-5	Sequence 5, Appl1	510	36.5	22.0	1118	1	US-08-278-089A-2	Sequence 2, Appl1
438	36.5	22.0	831	1	US-08-254-359A-5	Sequence 5, Appl1	511	36.5	22.0	1118	2	US-08-838-957A-2	Sequence 2, Appl1
439	36.5	22.0	831	1	US-08-483-043-5	Sequence 5, Appl1	512	36.5	22.0	1122	1	US-08-278-089A-6	Sequence 6, Appl1
440	36.5	22.0	831	1	US-08-481-238-5	Sequence 5, Appl1	513	36.5	22.0	1122	2	US-08-838-957A-6	Sequence 6, Appl1
441	36.5	22.0	831	2	US-08-471-066B-5	Sequence 5, Appl1	514	36.5	22.0	1124	1	US-08-322-474-2	Sequence 2, Appl1
442	36.5	22.0	831	2	US-08-484-956-5	Sequence 5, Appl1	515	36.5	22.0	1124	4	PCT-US93-06093-2	Sequence 2, Appl1
443	36.5	22.0	831	2	US-08-757-653-5	Sequence 5, Appl1	516	36.5	22.0	1205	4	US-09-352-159-29	Sequence 29, Appl1
444	36.5	22.0	831	2	US-08-599-491-5	Sequence 5, Appl1	517	36.5	22.0	1205	4	US-09-352-159-29	Sequence 29, Appl1
445	36.5	22.0	831	2	US-08-756-386-5	Sequence 5, Appl1	518	36.5	22.0	15	2	US-08-433-871C-73	Sequence 73, Appl1
446	36.5	22.0	831	2	US-08-823-516-5	Sequence 5, Appl1	519	36.5	22.0	15	4	US-07-270-956-73	Sequence 73, Appl1
447	36.5	22.0	831	3	US-08-682-853A-5	Sequence 5, Appl1	520	36.5	22.0	30	1	US-07-828-945A-3	Sequence 3, Appl1
448	36.5	22.0	831	3	US-08-759-038-5	Sequence 5, Appl1	521	36.5	22.0	63	1	US-08-447-411-61	Sequence 61, Appl1
449	36.5	22.0	831	3	US-08-758-314-5	Sequence 5, Appl1	522	36.5	22.0	63	2	US-08-662-227-18	Sequence 18, Appl1
450	36.5	22.0	831	4	US-09-350-309-5	Sequence 5, Appl1	523	36.5	22.0	63	4	US-09-017-947-18	Sequence 18, Appl1
451	36.5	22.0	831	4	US-08-520-946-5	Sequence 5, Appl1	524	36.5	22.0	152	4	US-08-936-165A-484	Sequence 484, App
452	36.5	22.0	832	1	US-07-977-434-2	Sequence 2, Appl1	525	36.5	22.0	186	1	US-08-117-083-22	Sequence 22, Appl1
453	36.5	22.0	832	1	US-08-156-020-2	Sequence 2, Appl1	526	36.5	22.0	201	2	US-08-688-342-1	Sequence 1, Appl1
454	36.5	22.0	832	1	US-08-156-020-6	Sequence 6, Appl1	527	36.5	22.0	201	2	US-09-113-788-1	Sequence 1, Appl1
455	36.5	22.0	832	1	US-08-156-020-4	Sequence 4, Appl1	528	36.5	22.0	231	2	US-08-902-655A-6	Sequence 6, Appl1
456	36.5	22.0	832	1	US-08-156-020-8	Sequence 8, Appl1	529	36.5	22.0	241	4	US-09-112-488A-10	Sequence 10, Appl1
457	36.5	22.0	832	1	US-08-156-020-10	Sequence 10, Appl1	530	36.5	22.0	258	3	US-09-035-706-3	Sequence 3, Appl1
458	36.5	22.0	832	1	US-08-073-384C-4	Sequence 4, Appl1	531	36.5	22.0	258	4	US-08-955-841-3	Sequence 3, Appl1
459	36.5	22.0	832	1	US-08-254-359A-4	Sequence 4, Appl1	532	36.5	22.0	258	4	US-09-390-425-3	Sequence 3, Appl1
460	36.5	22.0	832	1	US-08-483-043-4	Sequence 4, Appl1	533	36.5	22.0	258	4	US-09-556-906-3	Sequence 3, Appl1
461	36.5	22.0	832	1	US-08-458-819-2	Sequence 2, Appl1	534	36.5	22.0	269	2	US-08-701-191A-35	Sequence 35, Appl1
462	36.5	22.0	832	1	US-08-481-238-4	Sequence 4, Appl1	535	36.5	22.0	276	2	US-07-857-224B-72	Sequence 72, Appl1
463	36.5	22.0	832	2	US-08-471-066B-4	Sequence 4, Appl1	536	36.5	22.0	277	4	US-08-507-431-6	Sequence 558, App
464	36.5	22.0	832	2	US-08-484-956-4	Sequence 4, Appl1	537	36.5	22.0	296	1	US-08-507-431-6	Sequence 6, Appl1
465	36.5	22.0	832	2	US-08-757-653-4	Sequence 4, Appl1	538	36.5	22.0	296	3	US-09-116-622-6	Sequence 6, Appl1

539	36	21.7	295	4	US-09-219-277-6	Sequence 6, Appl1	612	36	21.7	745	2	US-09-099-124A-4	Sequence 4, Appl1
540	36	21.7	296	4	US-09-359-661-6	Sequence 6, Appl1	613	36	21.7	745	4	US-09-032-476-4	Sequence 4, Appl1
541	36	21.7	298	2	US-08-061-636-3	Sequence 3, Appl1	614	36	21.7	745	4	US-08-890-854-4	Sequence 4, Appl1
542	36	21.7	298	2	US-08-874-347-19	Sequence 19, Appl1	615	36	21.7	745	4	US-09-023-324-3	Sequence 4, Appl1
543	36	21.7	298	3	US-09-093-542-19	Sequence 19, Appl1	616	36	21.7	745	4	US-09-168-629-2	Sequence 2, Appl1
544	36	21.7	298	4	US-09-427-261-1	Sequence 1, Appl1	617	36	21.7	745	4	US-08-810-820-10	Sequence 10, Appl1
545	36	21.7	298	5	PCT-US94-05268-3	Sequence 3, Appl1	618	36	21.7	745	4	US-08-810-131A-2	Sequence 2, Appl1
546	36	21.7	312	1	US-08-915-003-2	Sequence 2, Appl1	619	36	21.7	745	3	US-08-909-954-4	Sequence 4, Appl1
547	36	21.7	312	2	US-08-642-247-2	Sequence 2, Appl1	620	36	21.7	865	4	US-09-612-204B-24	Sequence 24, Appl1
548	36	21.7	319	3	US-09-130-749-2	Sequence 2, Appl1	621	36	21.7	933	1	US-08-444-792-2	Sequence 2, Appl1
549	36	21.7	319	3	US-09-130-749-2	Sequence 2, Appl1	622	36	21.7	993	1	US-08-445-042-2	Sequence 2, Appl1
550	36	21.7	336	4	US-09-131-648-1	Sequence 21, Appl1	623	36	21.7	1036	2	US-08-542-003-6	Sequence 6, Appl1
551	36	21.7	336	4	US-09-457-040B-21	Sequence 21, Appl1	624	36	21.7	1036	2	US-08-222-760A-6	Sequence 6, Appl1
552	36	21.7	337	4	US-09-134-001C-5550	Sequence 550, Ap	625	36	21.7	1036	2	US-09-316-949-6	Sequence 6, Appl1
553	36	21.7	339	4	US-08-171-461-6	Sequence 35, Appl1	626	36	21.7	1039	4	US-09-409-648-8	Sequence 8, Appl1
554	36	21.7	335	4	US-09-134-001C-4718	Sequence 4718, Ap	627	36	21.7	1039	4	US-09-409-648-8	Sequence 8, Appl1
555	36	21.7	367	4	US-09-390-326-9	Sequence 9, Appl1	628	36	21.7	1039	4	US-09-409-648-8	Sequence 8, Appl1
556	36	21.7	373	3	US-08-746-883-4	Sequence 4, Appl1	629	36	21.7	1039	6	5196511-2	Patent No. 5196511
557	36	21.7	375	2	US-08-506-864A-1	Sequence 1, Appl1	630	36	21.7	1088	2	US-08-742-026-2	Sequence 2, Appl1
558	36	21.7	380	3	US-08-150-133-9	Sequence 1, Appl1	631	36	21.7	1088	2	US-08-742-026-23	Sequence 23, Appl1
559	36	21.7	380	3	US-08-150-133-9	Sequence 9, Appl1	632	36	21.7	1088	2	US-08-680-326-39	Sequence 39, Appl1
560	36	21.7	380	3	US-08-150-141-9	Sequence 9, Appl1	633	36	21.7	1037	2	US-08-340-011-5	Sequence 5, Appl1
561	36	21.7	380	4	US-09-374-493-9	Sequence 9, Appl1	634	36	21.7	1311	3	US-08-901-710-5	Sequence 5, Appl1
562	36	21.7	380	4	US-09-374-493-9	Sequence 9, Appl1	635	36	21.7	1338	3	US-08-750-141A-3	Sequence 3, Appl1
563	36	21.7	380	4	US-09-374-492-9	Sequence 9, Appl1	636	36	21.7	1350	2	US-08-319-866-9	Sequence 9, Appl1
564	36	21.7	383	4	US-09-134-001C-3429	Sequence 3429, Ap	637	36	21.7	1388	2	US-08-685-576-1	Sequence 1, Appl1
565	36	21.7	392	1	US-08-723-938-1	Sequence 1, Appl1	638	36	21.7	1388	2	US-08-685-576-1	Sequence 1, Appl1
566	36	21.7	392	2	US-09-080-538-1	Sequence 2, Appl1	639	36	21.7	1642	1	US-08-447-411-45	Sequence 45, Appl1
567	36	21.7	397	4	US-08-386-888A-2	Sequence 5, Appl1	640	36	21.7	1642	2	US-08-662-227-2	Sequence 2, Appl1
568	36	21.7	401	2	US-08-549-004A-5	Sequence 5, Appl1	641	36	21.7	1642	2	US-09-017-947-2	Sequence 2, Appl1
569	36	21.7	401	4	US-09-051-982A-5	Sequence 5, Appl1	642	36	21.7	1648	2	US-08-662-227-35	Sequence 35, Appl1
570	36	21.7	412	4	US-09-134-001C-3949	Sequence 3949, Ap	643	36	21.7	1648	4	US-09-017-947-35	Sequence 35, Appl1
571	36	21.7	418	2	US-08-494-907-18	Sequence 18, Appl1	644	36	21.7	1751	4	US-09-136-574A-44	Sequence 44, Appl1
572	36	21.7	418	5	PCT-US96-10986-18	Sequence 18, Appl1	645	36	21.7	2109	4	US-08-646-695-6	Sequence 6, Appl1
573	36	21.7	422	3	US-09-109-204-30	Sequence 30, Appl1	646	36	21.7	2109	5	PCT-US96-06053-6	Sequence 6, Appl1
574	36	21.7	445	2	US-09-047-026A-6	Sequence 6, Appl1	647	36	21.7	2152	4	US-09-036-987A-3	Sequence 3, Appl1
575	36	21.7	450	5	US-08-426-509A-7	Sequence 7, Appl1	648	36	21.7	2455	4	US-09-370-700-3	Sequence 48, Appl1
576	36	21.7	450	5	PCT-US95-05008-7	Sequence 7, Appl1	649	36	21.7	2455	4	US-09-413-814-48	Sequence 387, App
577	36	21.7	497	3	US-08-898-977-2	Sequence 2, Appl1	650	36	21.7	2972	2	US-08-469-260A-38	Sequence 24, Appl1
578	36	21.7	497	3	US-08-898-977-2	Sequence 2, Appl1	651	36	21.7	87	4	US-08-662-114B-24	Sequence 24, Appl1
579	36	21.7	498	1	US-09-535-171-2	Sequence 9, Appl1	652	35.5	21.4	87	4	US-08-729-594A-24	Sequence 24, Appl1
580	36	21.7	498	1	US-08-357-598-9	Sequence 9, Appl1	653	35.5	21.4	87	4	US-08-337-993-20	Sequence 10, Appl1
581	36	21.7	499	5	US-09-003-389-9	Sequence 9, Appl1	654	35.5	21.4	88	2	US-09-393-627B-16	Sequence 16, Appl1
582	36	21.7	517	4	PCT-US95-16435-9	Sequence 9, Appl1	655	35.5	21.4	179	4	US-09-393-627B-16	Sequence 1, Appl1
583	36	21.7	525	4	US-08-965-943-3	Sequence 3, Appl1	656	35.5	21.4	187	4	US-09-088-549-1	Sequence 20, Appl1
584	36	21.7	550	3	US-09-134-001C-3514	Sequence 3514, Ap	657	35.5	21.4	237	4	US-09-357-251-20	Sequence 10, Appl1
585	36	21.7	554	4	US-09-039-859-9	Sequence 9, Appl1	658	35.5	21.4	258	1	US-08-336-708A-10	Sequence 10, Appl1
586	36	21.7	568	4	US-09-222-938A-22	Sequence 22, Appl1	659	35.5	21.4	259	2	US-08-861-269-1	Sequence 1, Appl1
587	36	21.7	582	2	US-08-939-309-2	Sequence 20, Appl1	660	35.5	21.4	259	2	US-08-861-269-1	Sequence 1, Appl1
588	36	21.7	582	3	US-08-403-852D-20	Sequence 21, Appl1	661	35.5	21.4	259	2	US-09-134-596-1	Sequence 1, Appl1
589	36	21.7	582	4	US-08-510-646B-21	Sequence 20, Appl1	662	35.5	21.4	263	2	US-09-293-273-1	Sequence 1, Appl1
590	36	21.7	584	3	US-09-231-818-20	Sequence 1, Appl1	663	35.5	21.4	311	2	US-08-391-916A-6	Sequence 6, Appl1
591	36	21.7	590	4	US-09-160-483-1	Sequence 25, Appl1	664	35.5	21.4	327	3	US-08-391-916A-6	Sequence 5, Appl1
592	36	21.7	590	4	US-09-232-191-25	Sequence 75, Appl1	665	35.5	21.4	327	3	US-09-173-915-5	Sequence 5, Appl1
593	36	21.7	590	4	US-09-232-200-75	Sequence 75, Appl1	666	35.5	21.4	330	1	US-09-420-915-5	Sequence 2, Appl1
594	36	21.7	590	4	US-09-232-197-75	Sequence 75, Appl1	667	35.5	21.4	330	1	US-08-238-163-2	Sequence 2, Appl1
595	36	21.7	605	3	US-08-683-214-8	Sequence 8, Appl1	668	35.5	21.4	335	4	US-09-797-976-2	Sequence 2, Appl1
596	36	21.7	605	4	US-09-394-645-2	Sequence 2, Appl1	669	35.5	21.4	380	4	US-08-609-572-4	Sequence 4, Appl1
597	36	21.7	605	4	US-09-243-560B-2	Sequence 11, Appl1	670	35.5	21.4	380	4	US-08-841-751-4	Sequence 4, Appl1
598	36	21.7	612	2	US-08-673-789-11	Sequence 19, Appl1	671	35.5	21.4	380	4	US-08-846-340-4	Sequence 4, Appl1
599	36	21.7	629	4	US-09-300-909-19	Sequence 5, Appl1	672	35.5	21.4	383	2	US-08-846-344-4	Sequence 4, Appl1
600	36	21.7	630	3	US-08-394-177-5	Sequence 5, Appl1	673	35.5	21.4	398	1	US-08-507-431-2	Sequence 2, Appl1
601	36	21.7	630	3	US-08-394-912A-5	Sequence 5, Appl1	674	35.5	21.4	398	2	US-08-902-6655A-2	Sequence 2, Appl1
602	36	21.7	630	4	US-09-333-336-5	Sequence 5, Appl1	675	35.5	21.4	398	4	US-09-116-622-2	Sequence 2, Appl1
603	36	21.7	630	4	US-09-295-086-18	Sequence 18, Appl1	676	35.5	21.4	398	4	US-09-599-661-2	Sequence 2, Appl1
604	36	21.7	639	4	US-09-134-001C-3891	Sequence 3891, Ap	677	35.5	21.4	404	1	US-08-203-716-2	Sequence 2, Appl1
605	36	21.7	683	4	US-09-213-293D-1	Sequence 1, Appl1	678	35.5	21.4	404	1	US-08-442-663A-2	Sequence 2, Appl1
606	36	21.7	717	4	US-09-307-143-3	Sequence 3, Appl1	679	35.5	21.4	404	1	US-08-440-179-2	Sequence 2, Appl1
607	36	21.7	745	2	US-08-887-518-3	Sequence 3, Appl1	680	35.5	21.4	404	2	US-08-450-130A-1	Sequence 1, Appl1
608	36	21.7	745	2	US-09-023-121-3	Sequence 4, Appl1	681	35.5	21.4	404	2	US-08-391-916A-2	Sequence 2, Appl1
609	36	21.7	745	2	US-08-890-853-4	Sequence 3, Appl1	682	35.5	21.4	404	2	US-08-573-890-2	Sequence 2, Appl1
610	36	21.7	745	2	US-09-032-475-3	Sequence 4, Appl1	683	35.5	21.4	404	2	US-08-573-890-2	Sequence 2, Appl1
611	36	21.7	745	2	US-09-099-125A-4	Sequence 4, Appl1	684	35.5	21.4	404	3	US-08-450-362A-1	Sequence 1, Appl1

685	35.5	21.4	404.3	US-08-258-287B-39	Sequence 39, Appl	758	35	21.1	286.4	US-09-232-191-13	Sequence 13, Appl
686	35.5	21.4	404.3	US-08-368-704C-39	Sequence 39, Appl	759	35	21.1	286.4	US-09-232-200-13	Sequence 13, Appl
687	35.5	21.4	404.3	US-08-954-536-18	Sequence 18, Appl	760	35	21.1	286.4	US-09-232-197-13	Sequence 13, Appl
688	35.5	21.4	404.4	US-09-039-657-2	Sequence 2, Appl	761	35	21.1	286.4	US-09-232-201-13	Sequence 13, Appl
689	35.5	21.4	404.4	US-08-748-547-2	Sequence 2, Appl	762	35	21.1	286.4	US-08-858-207A-331	Sequence 331, Appl
690	35.5	21.4	404.4	US-08-908-436-3	Sequence 3, Appl	763	35	21.1	328.4	US-09-080-005-4	Sequence 4, Appl
691	35.5	21.4	404.4	US-09-248-179-2	Sequence 2, Appl	764	35	21.1	332.4	US-09-134-001C-3199	Sequence 3199, Appl
692	35.5	21.4	404.4	US-09-069-023-30	Sequence 30, Appl	765	35	21.1	344.2	US-08-690-493-1	Sequence 1, Appl
693	35.5	21.4	404.4	US-09-561-756-6	Sequence 6, Appl	766	35	21.1	353.1	US-08-687-702-37	Sequence 37, Appl
694	35.5	21.4	404.4	US-07-891-942G-10	Sequence 10, Appl	767	35	21.1	368.2	US-08-378-617A-11	Sequence 11, Appl
695	35.5	21.4	404.4	US-08-983-502-13	Sequence 13, Appl	768	35	21.1	387.4	US-09-134-001C-5036	Sequence 5036, Appl
696	35.5	21.4	404.5	PCR-US93-05705-4	Sequence 4, Appl	769	35	21.1	400.4	US-09-150-347-3	Sequence 3, Appl
697	35.5	21.4	404.5	PCR-US95-06132-2	Sequence 2, Appl	770	35	21.1	410.3	US-09-083-321-3	Sequence 3, Appl
698	35.5	21.4	404.5	PCR-US95-07619-1	Sequence 1, Appl	771	35	21.1	432.3	US-09-212-609B-23	Sequence 23, Appl
699	35.5	21.4	404.5	PCR-US96-10521-13	Sequence 13, Appl	772	35	21.1	433.4	US-09-212-609B-9	Sequence 9, Appl
700	35.5	21.4	404.5	US-08-985-343-5	Sequence 5, Appl	773	35	21.1	433.4	US-09-212-609B-18	Sequence 18, Appl
701	35.5	21.4	508.1	US-07-891-942G-10	Sequence 10, Appl	774	35	21.1	434.1	US-07-952-817-9	Sequence 9, Appl
702	35.5	21.4	529.1	US-07-891-942G-8	Sequence 8, Appl	775	35	21.1	434.1	US-07-952-817-14	Sequence 14, Appl
703	35.5	21.4	529.2	US-08-370-909-19	Sequence 19, Appl	776	35	21.1	434.1	US-07-952-817-14	Sequence 14, Appl
704	35.5	21.4	529.2	US-08-504-048-8	Sequence 8, Appl	777	35	21.1	434.1	5210025-7	Patent No. 5210025
705	35.5	21.4	535.4	US-09-117-860-24	Sequence 24, Appl	778	35	21.1	434.1	US-09-212-609B-15	Sequence 15, Appl
706	35.5	21.4	537.4	US-08-540-922D-12	Sequence 12, Appl	779	35	21.1	488.4	US-09-212-609B-20	Sequence 20, Appl
707	35.5	21.4	560.1	US-07-891-942G-5	Sequence 5, Appl	780	35	21.1	495.1	US-07-841-997A-2	Sequence 2, Appl
708	35.5	21.4	606.2	US-08-883-534-3	Sequence 3, Appl	781	35	21.1	495.1	US-08-280-301-2	Sequence 2, Appl
709	35.5	21.4	606.3	US-09-204-764-3	Sequence 3, Appl	782	35	21.1	495.1	US-08-588-983-2	Sequence 2, Appl
710	35.5	21.4	729.4	US-09-291-922-29	Sequence 29, Appl	783	35	21.1	495.1	US-08-588-976-2	Sequence 2, Appl
711	35.5	21.4	785.3	US-09-265-108-2	Sequence 2, Appl	784	35	21.1	495.1	US-05-013-98-2	Sequence 2, Appl
712	35.5	21.4	785.3	US-09-479-264-2	Sequence 2, Appl	785	35	21.1	495.1	US-08-985-343-2	Sequence 2, Appl
713	35.5	21.4	868.4	US-09-398-395A-56	Sequence 56, Appl	786	35	21.1	496.4	US-09-212-609B-6	Sequence 6, Appl
714	35.5	21.4	894.2	US-08-867-941-15	Sequence 15, Appl	787	35	21.1	496.4	US-09-032-365A-19	Sequence 19, Appl
715	35.5	21.4	1094.2	US-08-680-326-40	Sequence 40, Appl	788	35	21.1	543.4	US-09-739-455-4	Sequence 4, Appl
716	35.5	21.4	1802.1	US-08-232-537-2	Sequence 2, Appl	789	35	21.1	543.4	US-08-739-455-4	Sequence 4, Appl
717	35.5	21.4	1864.2	US-08-804-227C-3	Sequence 3, Appl	790	35	21.1	543.4	US-08-739-455-4	Sequence 4, Appl
718	35.5	21.4	2432.1	US-09-074-658-15	Sequence 15, Appl	791	35	21.1	544.4	US-08-732-025-4	Sequence 4, Appl
719	35.5	21.4	3033.1	US-07-925-695-8	Sequence 8, Appl	792	35	21.1	588.2	US-08-339-715A-1	Sequence 1, Appl
720	35.5	21.4	3033.1	US-07-925-695-9	Sequence 9, Appl	793	35	21.1	588.2	US-08-339-715A-3	Sequence 3, Appl
721	35.5	21.4	6095.2	US-09-144-085-2	Sequence 2, Appl	794	35	21.1	588.2	US-08-339-715A-4	Sequence 4, Appl
722	35.5	21.1	35.2	US-08-284-391B-47	Sequence 47, Appl	795	35	21.1	588.2	US-08-339-715A-5	Sequence 5, Appl
723	35.5	21.1	35.4	US-09-218-950-47	Sequence 47, Appl	796	35	21.1	615.1	US-08-638-911A-35	Sequence 35, Appl
724	35.5	21.1	68.1	US-08-466-093-148	Sequence 148, Appl	797	35	21.1	620.4	US-08-232-300-49	Sequence 49, Appl
725	35.5	21.1	68.2	US-08-444-733-148	Sequence 148, Appl	798	35	21.1	620.4	US-09-232-197-49	Sequence 49, Appl
726	35.5	21.1	68.2	US-08-464-134-148	Sequence 148, Appl	799	35	21.1	620.4	US-09-232-201-49	Sequence 49, Appl
727	35.5	21.1	68.2	US-08-461-361-148	Sequence 148, Appl	800	35	21.1	648.2	US-08-817-336A-2	Sequence 2, Appl
728	35.5	21.1	68.2	US-08-485-910-148	Sequence 148, Appl	801	35	21.1	699.4	US-08-274-121B-6	Sequence 6, Appl
729	35.5	21.1	107.2	US-08-993-228-16	Sequence 16, Appl	802	35	21.1	710.4	US-09-079-812E-2	Sequence 2, Appl
730	35.5	21.1	121.4	US-09-199-637A-421	Sequence 421, Appl	803	35	21.1	718.2	US-08-560-398-12	Sequence 12, Appl
731	35.5	21.1	139.4	US-08-542-363-31	Sequence 31, Appl	804	35	21.1	720.4	US-09-624-693A-13	Sequence 13, Appl
732	35.5	21.1	139.4	US-09-100-085-31	Sequence 31, Appl	805	35	21.1	726.4	US-08-624-693A-21	Sequence 21, Appl
733	35.5	21.1	139.4	US-09-134-001C-4916	Sequence 4916, Appl	806	35	21.1	830.3	US-08-872-855-11	Sequence 11, Appl
734	35.5	21.1	139.4	US-09-670-827-31	Sequence 31, Appl	807	35	21.1	832.4	US-08-981-492-6	Sequence 6, Appl
735	35.5	21.1	140.1	US-07-661-610C-10	Sequence 10, Appl	808	35	21.1	833.1	US-08-264-534-6	Sequence 6, Appl
736	35.5	21.1	143.4	US-09-605-785-482	Sequence 482, Appl	809	35	21.1	833.1	US-08-083-590A-2	Sequence 2, Appl
737	35.5	21.1	143.4	US-09-439-313-482	Sequence 482, Appl	810	35	21.1	833.1	US-08-465-500-6	Sequence 6, Appl
738	35.5	21.1	188.2	US-08-903-267-2	Sequence 2, Appl	811	35	21.1	833.2	US-08-346-126-6	Sequence 6, Appl
739	35.5	21.1	188.2	US-09-314-051-2	Sequence 2, Appl	812	35	21.1	833.2	US-08-346-126-6	Sequence 6, Appl
740	35.5	21.1	194.4	US-09-218-363-19	Sequence 19, Appl	813	35	21.1	833.3	US-08-532-84-2	Sequence 2, Appl
741	35.5	21.1	203.1	US-08-264-534-3	Sequence 3, Appl	814	35	21.1	833.3	US-08-893-828-6	Sequence 6, Appl
742	35.5	21.1	203.1	US-08-465-500-3	Sequence 3, Appl	815	35	21.1	833.3	US-08-216-260-4	Sequence 4, Appl
743	35.5	21.1	203.2	US-08-346-126-3	Sequence 3, Appl	816	35	21.1	853.1	US-08-638-911A-25	Sequence 25, Appl
744	35.5	21.1	203.2	US-08-346-126-3	Sequence 3, Appl	817	35	21.1	853.1	US-08-638-911A-25	Sequence 25, Appl
745	35.5	21.1	203.2	US-08-893-828-3	Sequence 3, Appl	818	35	21.1	853.1	US-08-638-911A-29	Sequence 29, Appl
746	35.5	21.1	208.4	US-09-134-001C-3785	Sequence 3785, Appl	819	35	21.1	853.1	US-08-638-911A-31	Sequence 31, Appl
747	35.5	21.1	222.2	US-09-384-162-7	Sequence 7, Appl	820	35	21.1	866.2	US-08-456-461A-6	Sequence 6, Appl
748	35.5	21.1	231.2	US-08-622-354-5	Sequence 5, Appl	821	35	21.1	874.2	US-08-456-461A-6	Sequence 6, Appl
749	35.5	21.1	236.1	US-08-400-159-4	Sequence 4, Appl	822	35	21.1	874.2	US-08-237-401A-6	Sequence 6, Appl
750	35.5	21.1	236.3	US-08-611-729A-4	Sequence 4, Appl	823	35	21.1	875.4	US-09-134-001C-3943	Sequence 3943, Appl
751	35.5	21.1	239.4	US-08-634-475-4	Sequence 4, Appl	824	35	21.1	880.3	US-08-445-640-10	Sequence 10, Appl
752	35.5	21.1	239.4	US-09-709-791-4	Sequence 4, Appl	825	35	21.1	880.3	US-08-170-558-10	Sequence 10, Appl
753	35.5	21.1	255.1	US-07-795-859B-2	Sequence 2, Appl	826	35	21.1	880.3	US-08-447-314-10	Sequence 10, Appl
754	35.5	21.1	255.1	US-08-457-616-2	Sequence 2, Appl	827	35	21.1	880.3	US-08-445-461-10	Sequence 10, Appl
755	35.5	21.1	266.5	US-08-443-568B-16	Sequence 16, Appl	828	35	21.1	949.3	US-09-404-627-4	Sequence 4, Appl
756	35.5	21.1	266.5	PCR-US94-06997-16	Sequence 16, Appl	829	35	21.1	1013.4	US-08-860-886-2	Sequence 2, Appl
757	35.5	21.1	268.4	US-09-218-363-12	Sequence 12, Appl	830	35	21.1	1014.4	US-09-078-347A-3	Sequence 3, Appl

831	35	21.1	1205	2	US-08-319-866-10	Sequence 10, Appl	904	34.5	20.8	614	5	PCT-US95-10579-2	Sequence 2, Appl
832	35	21.1	1205	4	US-09-123-708-6	Sequence 6, Appl	905	34.5	20.8	626	1	US-08-472-934-6	Sequence 6, Appl
833	35	21.1	1205	4	US-09-123-624-6	Sequence 6, Appl	906	34.5	20.8	626	2	US-08-323-460A-6	Sequence 6, Appl
834	35	21.1	1247	4	US-09-404-627-4	Sequence 2, Appl	907	34.5	20.8	626	2	US-08-461-146C-6	Sequence 6, Appl
835	35	21.1	1529	4	US-09-134-001C-3945	Sequence 3945, Ap	908	34.5	20.8	626	3	US-08-461-145C-6	Sequence 6, Appl
836	35	21.1	1693	3	US-08-478-507-7	Sequence 7, Appl	909	34.5	20.8	626	4	US-09-423-890-6	Sequence 6, Appl
837	35	21.1	1693	3	US-08-128-753A-7	Sequence 7, Appl	910	34.5	20.8	626	4	US-09-423-890-12	Sequence 12, Appl
838	35	21.1	1693	4	US-09-553-427-7	Sequence 7, Appl	911	34.5	20.8	626	4	US-08-628-829-10	Sequence 10, Appl
839	35	21.1	1781	2	US-08-477-451-11	Sequence 11, Appl	912	34.5	20.8	666	4	US-09-134-001C-4739	Sequence 4739, Ap
840	35	21.1	2210	4	US-09-309-572-2	Sequence 7, Appl	913	34.5	20.8	691	5	PCT-US91-08442-2	Sequence 2, Appl
841	35	21.1	2476	2	US-08-276-967-2	Sequence 2, Appl	914	34.5	20.8	725	2	US-08-813-940-25	Sequence 25, Appl
842	35	21.1	2710	2	US-08-568-459A-12	Sequence 12, Appl	915	34.5	20.8	752	4	US-09-817-180-2	Sequence 2, Appl
843	35	21.1	2710	2	US-08-487-826B-12	Sequence 12, Appl	916	34.5	20.8	780	1	US-08-188-228-50	Sequence 50, Appl
844	35	21.1	2710	4	US-09-210-288-12	Sequence 12, Appl	917	34.5	20.8	780	1	US-08-332-643-44	Sequence 44, Appl
845	35	21.1	2873	1	US-08-466-033-15	Sequence 15, Appl	918	34.5	20.8	780	1	US-08-332-638-50	Sequence 50, Appl
846	35	21.1	2873	1	US-08-638-911A-2	Sequence 2, Appl	919	34.5	20.8	822	4	US-09-817-180-4	Sequence 4, Appl
847	35	21.1	2873	2	US-08-444-733-15	Sequence 15, Appl	920	34.5	20.8	888	1	US-08-445-640-35	Sequence 35, Appl
848	35	21.1	2873	2	US-08-464-134-15	Sequence 15, Appl	921	34.5	20.8	888	3	US-08-170-558-35	Sequence 35, Appl
849	35	21.1	2873	2	US-08-461-361-15	Sequence 15, Appl	922	34.5	20.8	888	3	US-08-447-314-35	Sequence 35, Appl
850	35	21.1	2873	2	US-08-485-910-15	Sequence 15, Appl	923	34.5	20.8	888	3	US-08-445-461-35	Sequence 35, Appl
851	35	21.1	2873	5	PCT-US95-06266-15	Sequence 15, Appl	924	34.5	20.8	888	2	US-08-867-941-11	Sequence 11, Appl
852	35	21.1	2910	1	US-08-466-033-183	Sequence 183, App	925	34.5	20.8	1087	2	US-08-570-311-8	Sequence 8, Appl
853	35	21.1	2910	2	US-08-444-733-183	Sequence 183, App	926	34.5	20.8	1087	2	US-08-353-485-8	Sequence 8, Appl
854	35	21.1	2910	2	US-08-464-134-183	Sequence 183, App	927	34.5	20.8	1135	2	US-08-469-537A-97	Sequence 97, Appl
855	35	21.1	2910	2	US-08-461-361-183	Sequence 183, App	928	34.5	20.8	1135	2	US-08-323-474-8	Sequence 8, Appl
856	35	21.1	2910	2	US-08-485-910-183	Sequence 183, App	929	34.5	20.8	1138	2	US-08-469-537A-98	Sequence 98, Appl
857	35	21.1	2910	5	PCT-US95-06266-157	Sequence 157, App	930	34.5	20.8	1138	2	US-08-320-240A-5	Sequence 5, Appl
858	35	21.1	3060	2	US-08-467-826B-14	Sequence 14, Appl	931	34.5	20.8	1285	2	US-08-540-406-6	Sequence 6, Appl
859	34.5	20.8	133	4	US-09-072-596-783	Sequence 783, App	932	34.5	20.8	1285	3	US-08-556-085-6	Sequence 6, Appl
860	34.5	20.8	150	4	US-09-605-785-707	Sequence 707, App	933	34.5	20.8	1285	4	US-08-554-668-6	Sequence 6, Appl
861	34.5	20.8	184	4	US-09-552-322-11	Sequence 11, Appl	934	34.5	20.8	1285	4	US-08-918-658-6	Sequence 6, Appl
862	34.5	20.8	223	2	US-08-430-633-1	Sequence 1, Appl	935	34.5	20.8	1285	5	PCT-US95-13233-6	Sequence 6, Appl
863	34.5	20.8	223	2	US-08-936-854-1	Sequence 1, Appl	936	34.5	20.8	1286	4	US-09-268-110-3	Sequence 3, Appl
864	34.5	20.8	236	1	US-08-158-682A-4	Sequence 4, Appl	937	34.5	20.8	1289	4	US-08-674-509B-48	Sequence 48, Appl
865	34.5	20.8	236	1	US-08-687-695-4	Sequence 4, Appl	938	34.5	20.8	1289	4	US-08-954-698-48	Sequence 48, Appl
866	34.5	20.8	236	1	US-08-816-241-4	Sequence 4, Appl	939	34.5	20.8	1358	2	US-08-570-311-27	Sequence 27, Appl
867	34.5	20.8	236	2	US-09-040-482-4	Sequence 4, Appl	940	34.5	20.8	1487	3	US-08-840-062-7	Sequence 7, Appl
868	34.5	20.8	236	2	US-09-128-395-4	Sequence 4, Appl	941	34.5	20.8	1687	2	US-08-570-311-29	Sequence 29, Appl
869	34.5	20.8	262	2	US-07-857-224B-62	Sequence 62, Appl	942	34.5	20.8	1704	3	US-08-336-308A-10	Sequence 10, Appl
870	34.5	20.8	272	2	US-07-857-224B-63	Sequence 63, Appl	943	34.5	20.8	1704	3	US-08-822-324-6	Sequence 6, Appl
871	34.5	20.8	280	4	US-09-362-473-8	Sequence 8, Appl	944	34.5	20.8	1704	4	US-09-490-931-10	Sequence 10, Appl
872	34.5	20.8	290	1	US-08-440-846-2	Sequence 2, Appl	945	34.5	20.8	1732	2	US-08-570-311-10	Sequence 10, Appl
873	34.5	20.8	291	2	US-08-701-191A-22	Sequence 22, Appl	946	34.5	20.8	1732	2	US-08-353-485-10	Sequence 10, Appl
874	34.5	20.8	291	2	US-08-701-191A-22	Sequence 22, Appl	947	34.5	20.8	2206	1	US-07-852-260-2	Sequence 2, Appl
875	34.5	20.8	305	4	US-08-858-207A-392	Sequence 392, App	948	34.5	20.8	2206	2	US-08-461-503-2	Sequence 2, Appl
876	34.5	20.8	308	2	US-08-576-626A-36	Sequence 36, Appl	949	34.5	20.8	2206	4	US-08-465-250-12	Sequence 12, Appl
877	34.5	20.8	348	4	US-09-134-001C-2874	Sequence 2874, Ap	950	34.5	20.8	2439	4	US-09-074-658-11	Sequence 11, Appl
878	34.5	20.8	376	4	US-09-125-089-4	Sequence 4, Appl	951	34.5	20.8	2930	4	US-09-417-822-2	Sequence 2, Appl
879	34.5	20.8	376	4	US-09-125-089-6	Sequence 8, Appl	952	34.5	20.8	3177	2	US-08-477-851-4	Sequence 4, Appl
880	34.5	20.8	376	4	US-09-125-089-10	Sequence 10, Appl	953	34.5	20.8	3177	1	US-08-383-753-52	Sequence 52, Appl
881	34.5	20.8	376	4	US-09-134-001C-5654	Sequence 5654, Ap	954	34	20.5	18	2	US-08-586-772-52	Sequence 52, Appl
882	34.5	20.8	376	4	US-09-734-257-4	Sequence 4, Appl	955	34	20.5	18	2	US-08-959-512-52	Sequence 52, Appl
883	34.5	20.8	376	4	US-09-734-257-8	Sequence 8, Appl	956	34	20.5	18	2	US-09-512-983-52	Sequence 52, Appl
884	34.5	20.8	376	4	US-09-734-257-10	Sequence 10, Appl	957	34	20.5	18	4	US-08-469-260A-558	Sequence 558, App
885	34.5	20.8	376	4	US-09-734-257-12	Sequence 12, Appl	958	34	20.5	40	4	US-09-154-083-16	Sequence 16, Appl
886	34.5	20.8	411	4	US-09-188-930-302	Sequence 302, App	959	34	20.5	46	4	US-08-679-493A-169	Sequence 169, App
887	34.5	20.8	413	4	US-09-215-694-6	Sequence 6, Appl	960	34	20.5	84	4	US-08-679-493A-100	Sequence 100, App
888	34.5	20.8	425	4	US-09-134-001C-5654	Sequence 5654, Ap	961	34	20.5	111	1	US-08-466-886-36	Sequence 36, Appl
889	34.5	20.8	455	3	US-09-362-473-6	Sequence 6, Appl	962	34	20.5	111	4	US-08-469-617-36	Sequence 36, Appl
890	34.5	20.8	505	4	US-08-747-221B-14	Sequence 14, Appl	963	34	20.5	133	3	US-09-012-515A-18	Sequence 18, Appl
891	34.5	20.8	505	4	US-09-005-051-14	Sequence 14, Appl	964	34	20.5	133	3	US-08-360-144A-18	Sequence 18, Appl
892	34.5	20.8	513	4	US-09-097-889-15	Sequence 15, Appl	965	34	20.5	133	5	PCT-US95-06722-18	Sequence 18, Appl
893	34.5	20.8	535	4	US-09-117-860-18	Sequence 18, Appl	966	34	20.5	134	3	PCT-US95-06722-18	Sequence 18, Appl
894	34.5	20.8	535	4	US-09-134-001C-5244	Sequence 5244, Ap	967	34	20.5	134	3	US-08-806-597A-8	Sequence 8, Appl
895	34.5	20.8	550	3	US-08-747-221B-58	Sequence 58, Appl	968	34	20.5	145	5	US-08-970-428A-8	Sequence 28, Appl
896	34.5	20.8	552	4	US-09-005-051-58	Sequence 58, Appl	969	34	20.5	165	2	PCT-US94-10261A-28	Sequence 28, Appl
897	34.5	20.8	552	1	US-07-662-223-2	Sequence 2, Appl	970	34	20.5	165	2	US-08-777-321A-1	Sequence 1, Appl
898	34.5	20.8	576	3	US-08-948-564-16	Sequence 16, Appl	971	34	20.5	207	4	US-09-134-001C-4726	Sequence 4726, Ap
899	34.5	20.8	601	4	US-09-336-663A-4	Sequence 4, Appl	972	34	20.5	208	2	US-08-606-143-41	Sequence 41, Appl
900	34.5	20.8	612	1	US-08-344-635-2	Sequence 2, Appl	973	34	20.5	220	2	US-08-763-121-4	Sequence 4, Appl
901	34.5	20.8	614	1	US-08-543-881-2	Sequence 2, Appl	974	34	20.5	220	2	US-09-216-066-4	Sequence 4, Appl
902	34.5	20.8	614	1	US-08-291-299-2	Sequence 2, Appl	975	34	20.5	236	4	US-09-154-750A-80	Sequence 80, Appl
903	34.5	20.8	614	5	PCT-US94-00119-2	Sequence 2, Appl	976	34	20.5	236	4		

977 34 20.5 237 2 US-08-578-709-11 Sequence 11, Appl
978 34 20.5 240 3 US-08-908-332-9 Sequence 9, Appl
979 34 20.5 242 3 US-08-908-332-1 Sequence 1, Appl
980 34 20.5 247 4 US-09-370-838-112 Sequence 112, Appl
981 34 20.5 251 4 US-08-811-682-17 Sequence 17, Appl
982 34 20.5 261 4 US-09-134-003C-4193 Sequence 4193, Ap
983 34 20.5 262 4 US-09-008-271A-2 Sequence 2, Appl
984 34 20.5 269 4 US-09-724-864-47 Sequence 47, Appl
985 34 20.5 270 4 US-09-362-473-10 Sequence 10, Appl
986 34 20.5 284 2 US-09-055-095-1 Sequence 1, Appl
987 34 20.5 290 1 US-08-202-186-26 Sequence 26, Appl
988 34 20.5 310 2 US-08-701-191A-8 Sequence 8, Appl
989 34 20.5 310 2 US-09-047-026A-25 Sequence 25, Appl
990 34 20.5 315 3 US-09-135-639-4 Sequence 4, Appl
991 34 20.5 323 4 US-09-237-119-2 Sequence 2, Appl
992 34 20.5 343 2 US-08-599-171A-28 Sequence 28, Appl
993 34 20.5 343 2 US-08-646-590B-28 Sequence 28, Appl
994 34 20.5 343 3 US-09-069-226-28 Sequence 28, Appl
995 34 20.5 343 3 US-09-412-184-28 Sequence 28, Appl
996 34 20.5 347 1 US-08-461-244-3 Sequence 3, Appl
997 34 20.5 355 4 US-09-390-131-5 Sequence 5, Appl
998 34 20.5 360 1 US-08-450-393A-4 Sequence 4, Appl
999 34 20.5 360 3 US-09-043-627-2 Sequence 2, Appl
1000 34 20.5 360 4 US-08-446-669-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-045-284A-2
Sequence 2, Application US/09045284A
Patent No. 6265192
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107U51
CURRENT APPLICATION NUMBER: US/09/045,284A
CURRENT FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 100.0% Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SMRSGSFVQGLFGQHPDYFLMEPAWHV 29
DB 50 SMRSGSFVQGLFGQHPDYFLMEPAWHV 78

RESULT 2
US-09-190-911-1
Sequence 1, Application US/09190911
Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/09/190,911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045,284
EARLIER FILING DATE: 1998-03-20

NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-09-190-911-1

Query Match 100.0% Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVQGLFGQHPDYFLMEPAWHV 29
DB 50 SMRSGSFVQGLFGQHPDYFLMEPAWHV 78

RESULT 3
US-09-263-023-2
Sequence 2, Application US/09263023
Patent No. 6037159
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Reiji
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263,023
CURRENT FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-263-023-2

Query Match 74.7% Score 124; DB 3; Length 483;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SMRSGSFVQGLFGQHPDYFLMEPAWHV 29
DB 125 TWRSGSFVQGLFGQHPDYFLMEPAWHV 153

RESULT 4
US-09-471-867-2
Sequence 2, Application US/09471867
Patent No. 6455289
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Reiji
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471,867
CURRENT FILING DATE: 1999-12-23
EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-471-867-2

Query Match 74.7%; Score 124; DB 4; Length 483;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFQGHDPDYFLMEPMHV 29
Db 125 TWRSGSFFGELFNQNPVEFLYEPVHV 153

RESULT 5
US-09-263-023-4
Sequence 4, Application US/09263023
Patent No. 6037159
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Reiji
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263.023
CURRENT FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-263-023-4

Query Match 74.7%; Score 124; DB 3; Length 484;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFQGHDPDYFLMEPMHV 29
Db 126 TWRSGSFFGELFNQNPVEFLYEPVHV 154

RESULT 6
US-09-471-867-4
Sequence 4, Application US/09471867
Patent No. 6455289
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Reiji
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471.867
CURRENT FILING DATE: 1999-12-23

EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-471-867-4

Query Match 74.7%; Score 124; DB 4; Length 484;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFQGHDPDYFLMEPMHV 29
Db 126 TWRSGSFFGELFNQNPVEFLYEPVHV 154

RESULT 7
US-09-015-188-2
Sequence 2, Application US/09015188C
Patent No. 6393358
GENERAL INFORMATION:
APPLICANT: Williams, Kevin J
APPLICANT: Tabas, Ira
TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
FILE REFERENCE: JEFF-0231
CURRENT APPLICATION NUMBER: US/09/015.188C
CURRENT FILING DATE: 1998-01-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-015-188-2

Query Match 69.9%; Score 116; DB 4; Length 411;
Best Local Similarity 81.5%; Pred. No. 8.5e-10;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 RSGGSFVGQLFQGHDPDYFLMEPMHV 29
Db 70 RSGGSFVGQLFQGHDPDYFLMEPMHV 96

RESULT 8
US-08-655-878-2
Sequence 2, Application US/08655878
Patent No. 5827713
GENERAL INFORMATION:
APPLICANT: FUKUTA, MASAKAZU
APPLICANT: HABUCHI, OSAMI
TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS:
STREET:
CITY:
STATE:
COUNTRY:
ZIP:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,878
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 458
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-878-2

Query Match 60.8%; Score 101; DB 2; Length 458;
Best Local Similarity 59.3%; Pred. No. 1.8e-07;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFFVGLFGQHPDVFYLMPEAMHV 29
|:|||||:|:|:|||||:|:
Db 122 RTGSSFFGFFFNQGNIFYLPFLMHI 148

RESULT 9
US-08-899-514-2
Sequence 2, Application US/08899514
Patent No. 5910581
GENERAL INFORMATION:
APPLICANT: HABUCHI, OSAMI
APPLICANT: FUKUTA, MASAAZU
TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
TITLE OF INVENTION: SULFOTRANSFERASE ORIGINATING FROM HUMAN AND DNA CODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR, LLP
STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CALIFORNIA
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,514
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: DANIEL E ALTMAN
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: T09AM21.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714 760 0404
TELEFAX: 714 760 9502
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 479
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-899-514-2

Query Match 60.8%; Score 101; DB 2; Length 479;
Best Local Similarity 59.3%; Pred. No. 1.9e-07;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFFVGLFGQHPDVFYLMPEAMHV 29
|:|||||:|:|:|||||:|:
Db 142 RTGSSFFGFFFNQGNIFYLPFLMHI 168

RESULT 10
US-08-443-865-6
Sequence 6, Application US/08443865
Patent No. 5654146
GENERAL INFORMATION:
APPLICANT: BRAXTON, SCOTT M.
APPLICANT: DELEGANE, ANGELO M.
APPLICANT: DIEP, DINH
TITLE OF INVENTION: HUMAN ICE HOMOLOG
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3330 Hillview Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,865
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PP-0037 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-443-865-6

Query Match 32.2%; Score 53.5; DB 1; Length 372;
Best Local Similarity 41.7%; Pred. No. 2.3;
Matches 10; Conservative 5; Mismatches 6; Indels 3; Gaps 1;

QY 1 SWRS---GSSFFVGLFGQHPDVFY 21
||| |||:|:|:|:
Db 308 SWDRTRGSIIFGELMHMLPEIFF 331

RESULT 11
US-08-900-491-6
Sequence 6, Application US/08900491
Patent No. 5808001
GENERAL INFORMATION:
APPLICANT: BRAXTON, SCOTT M.
APPLICANT: DELEGANE, ANGELO M.
APPLICANT: DIEP, DINH
TITLE OF INVENTION: HUMAN ICE HOMOLOG
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:


```

Db      356 GNEFVSQLTGHRVYKILQDKWM 378
| : | | | | | : |
RESULT 13
US-08-737-825-6
; Sequence 6, Application US/08737825
; Patent No. 5871922
GENERAL INFORMATION:
APPLICANT: SALMOND, GEORGE PEACOCK COPELAND
APPLICANT: MCGOWAN, SIMON JAMES
APPLICANT: SEBALIHA, MOHAMED
APPLICANT: COX, ANTHONY RICHARD JOHN
APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY
APPLICANT: PORTER, LAUREN ELIZABETH
APPLICANT: BYCROFT, BARRE WALSHAM
APPLICANT: WILLIAMS, PAUL
APPLICANT: STEWART, GORDON SIDNEY ANDERSON BIRNIE
TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,825
FILING DATE: 03-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 1009-0105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 246345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Erwiniia carotovora
US-08-737-825-6

Query Match          28.0%; Score 46.5; DB 2; Length 376;
Best Local Similarity 36.7%; Pred. No. 27;
Matches 11; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

Oy      1 SMRSGSFVGQLFGCHPDVFY-LMEPAWHV 29
| : | : | | | | | : | | : : |
Db      298 SLRAQRWTWGDLXGLGDHIITSLQDTGRV 327

RESULT 14
US-09-134-001C-5331
; Sequence 5331, Application US/09134001C
; Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
```

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5331
LENGTH: 165
TYPE: PR1
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5331

Query Match 27.1%; Score 45; DB 4; Length 165;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 OHPDVEYLMPEAW 27
DB 59 QEPVFTAMEPSW 71

RESULT 15
US-08-551-356-4
Sequence 4, Application US/08551336
Patent No. 5830700
GENERAL INFORMATION:
APPLICANT: Iran1, Meher
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,356
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/349,762
FILING DATE:
APPLICATION NUMBER: US/07/998,271
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Holly, Julie A
REGISTRATION NUMBER: 33-246
REFERENCE/DOCKET NUMBER: 92-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-551-356-4

Query Match 27.1%; Score 45; DB 2; Length 643;
Best Local Similarity 52.9%; Pred. No. 83;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 SWRSGSFVGOLEFGQHP 17
DB 333 SWRSGSGTGTGNCNP 349

DB 333 SWRSGSGTGTGNCNP 349

RESULT 16
PCT-US93-12687-4
Sequence 2, Application PC/TUS9312687
GENERAL INFORMATION:
APPLICANT: Iran1, Meher H.
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12687
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,271
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 92-26PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12687-4

Query Match 27.1%; Score 45; DB 5; Length 643;
Best Local Similarity 55.2.9%; Pred. No. 83;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRSGSFVGOLEFGQHP 17
DB 333 SWRSGSGTGTGNCNP 349

RESULT 17
US-08-206-176-2
Sequence 2, Application US/08206176
Patent No. 5639940
GENERAL INFORMATION:
APPLICANT: Garner, Ian
APPLICANT: Dalrymple, Michael A
APPLICANT: Prunkard, Donna E
TITLE OF INVENTION: Production of fibrinogen in transgenic
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:

SEQUENCE CHARACTERISTICS:
LENGTH: 1336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12687-6

Query Match 27.1%; Score 45; DB 5; Length 1336;
Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRSGSFEVQLFGQHP 17
DB 1026 SWNSSGSGTGTGNGNP 1042

RESULT 25
US-09-357-251-22
Sequence 22, Application US/09357251
Patent No. 6271441
GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
APPLICANT: Farnou, Layo O.
APPLICANT: Orozco, Buddy
APPLICANT: Schwaber, James S.
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
FILE REFERENCE: BB-1193
CURRENT APPLICATION NUMBER: US/09/357,251
CURRENT FILING DATE: 1999-07-20
EARLIER APPLICATION NUMBER: 60/093,530
EARLIER FILING DATE: July 21, 1998
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO: 22
LENGTH: 123
TYPE: PRT
ORGANISM: Glycine max
US-09-357-251-22

Query Match 26.5%; Score 44; DB 4; Length 123;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 12; Conservative 2; Mismatches 8; Indels 8; Gaps 2;

QY 2 WRSGSFEVQLFGQHP---DVEYLMEPA 26
DB 63 WNFDTLFGQ---QHPARDHDTFLSEPA 89

RESULT 26
US-09-052-877-2
Sequence 2, Application US/09052877
Patent No. 6190912
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: BLK GENES, GENE PRODUCTS AND USES
TITLE OF INVENTION: THEREOF IN APOPTOSIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/052,877
FILING DATE: 31-MAR-1998
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 480140.433
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 662-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-052-877-2

Query Match 26.5%; Score 44; DB 4; Length 150;
Best Local Similarity 37.9%; Pred. No. 23;
Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 SWRSGSFEVQLFGQHPDYFLMEPAHV 29
DB 119 AWSPDQDPQLFQVYLLVFLIGAWYL 147

RESULT 27
US-08-740-223A-15
Sequence 15, Application US/08740223A
Patent No. 6265564
GENERAL INFORMATION:
APPLICANT: Davls, et al.
TITLE OF INVENTION: Expressed ligand - Vascular
TITLE OF INVENTION: Intercellular Signalling Molecule
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/740,223A
FILING DATE: 25-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/999
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Covert, Robert J.
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7721
TELEFAX: 914-345-7400
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: TTL2
LOCATION: 1...496
OTHER INFORMATION: mouse TIE-2 ligand 2
US-08-740-223A-15

Query Match 26.5%; Score 44; DB 4; Length 496;

Best Local Similarity 34.8%; Pred. No. 88;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSFFVGLFGQHPDVFYLMPEPAW 27
Db 356 GNEFISQITGQHRVYKIQKDW 378

RESULT 28

US-09-709-188-15
; Sequence 15, Application US/09709188
; Patent No. 6441137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-2
; CURRENT APPLICATION NUMBER: US/09/709,188
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-709-188-15

Query Match 26.5%; Score 44; DB 4; Length 496;
Best Local Similarity 34.8%; Pred. No. 88;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSFFVGLFGQHPDVFYLMPEPAW 27
Db 356 GNEFISQITGQHRVYKIQKDW 378

RESULT 29

US-08-293-814E-4
; Sequence 4, Application US/08295814E
; Patent No. 5658786
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E.
; APPLICANT: Borden, Laurence A.
; APPLICANT: Hartig, Paul R.
; APPLICANT: Weinstanek, Richard L.
; TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
; TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/295,814E
; FILING DATE: DECEMBER 19, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 4058-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
; LENGTH: 627 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-295-814E-4

Query Match 26.5%; Score 44; DB 1; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 2 WR-----GSSFFVGLFGQHPDVFYLMPEPAHV 29
Db 117 WRVCPLEFGIGYATQVIEAHNLVYIIILAWAI 150

RESULT 30

US-08-291-299-10
; Sequence 10, Application US/08291299
; Patent No. 5766848
; GENERAL INFORMATION:
; APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINE/GABA
; TITLE OF INVENTION: TRANSPORTER AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/291,299
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 42897-A/JPM/TEP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 627 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-291-299-10

Query Match 26.5%; Score 44; DB 1; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 2 WR-----GSSFFVGLFGQHPDVFYLMPEPAHV 29
Db 117 WRVCPLEFGIGYATQVIEAHNLVYIIILAWAI 150

RESULT 31

US-09-343-361-4
; Sequence 4, Application US/09343361
; Patent No. 6225115
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses

```

1  TITLE OF INVENTION: Theoeof
2  FILE REFERENCE: 40558-D
3  CURRENT APPLICATION NUMBER: US-09/343,361
4  CURRENT FILING DATE: 1999-06-30
5  NUMBER OF SEQ ID NOS: 26
6  SOFTWARE: PatentIn Ver. 2.0 - beta
7  SEQ ID NO. 4
8  LENGTH: 627
9  TYPE: PRT
10 ORGANISM: Rattus norvegicus
11 IS-09-343-361-4

```

```

Query Match      26.5%; Score 44; DB 4; Length 627;
Best Local Similarity 26.5%; Pred. No. 1.1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps
QY      2 WR-----SSSPVGGLEFGCHDHYEYIMPRNV 29
      ||      ||      ||      ||      ||
Db      117 WRKVCPLFEIGATGYVTEAHNLVYIIILALAI 150

```

```

RESULT 32
PCT-US93-01959-4
: Sequence 4, Application PC/TUS9301959
: GENERAL INFORMATION:
: APPLICANT: Smith, E. Kelli
: APPLICANT: Borden, A. Laurence
: APPLICANT: Hartig, R. Paul
: APPLICANT: Weinschenk, L. Richard
: TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
: TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10112
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/01959
: FILING DATE: 19930304
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 40558A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-977-9550
: TELEFAX: 212-664-0525
: TELEX: 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 627 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US93-01959-4

```

Query Match:	26.5%	Score 44	DB 5:	Length 627
Best Local Similarity	26.5%	Pred. No. 1.1e+02		
Matches	9	Conservative	7	Mismatches 12; Indels 6; Gaps 1
QY	2 WR-----SSSFVGGLESGHPDVEYLMPEAMHV	29		
Db	117 WRRVCPLEFGIGATGVIEAHILNLYIIITLLAMAI	150		

```

1  RESULT 33
2  PCT-US95-10579-10
3  Sequence 10 Application PC/TUS9510579
4  GENERAL INFORMATION:
5  APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
6  TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINE/GABA
7  TITLE OF INVENTION: TRANSPORTER AND USES THEREOF
8  NUMBER OF SEQUENCES: 10
9  CORRESPONDENCE ADDRESS:
10 ADDRESS: Cooper & Dunham LLP
11 STREET: 1185 Avenue of the Americas
12 CITY: New York
13 STATE: New York
14 COUNTRY: U.S.A.
15 ZIP: 10036
16
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.24
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: PCT/US95/10579
24 FILING DATE:
25
26 CLASSIFICATION:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: White, John P.
29 REGISTRATION NUMBER: 28,678
30 REFERENCE/DOCKET NUMBER: 42897-A-PCT/JPW/MAT
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (212) 278-0400
33 TELEFAX: (212) 391-0525
34 INFORMATION FOR SEQ ID NO: 10:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 627 amino acids
37 TYPE: amino acid
38 TOPOLOGY: linear
39
40 MOLECULE TYPE: protein
41
42 PCT-US95-10579-10

```

```

Query Match      26.5%;  Score 44;  DB 5;  Length 627;
Best Local Similarity 26.5%;  Pred. No. 1.1e+02;
Matches 9;  Conservative 7;  Mismatches 12;  Indels 6;  Gaps 1

      OY      2  WR-----SSSTFVGLFGQHDFVTLMEPAHHV 29
          ||      |      |      |      |      |
      Db      117  WRKVCPLFEGIGATVATLPAHLNLTIIILAAVI 150

```

RESULT 34
 US-08/295-814E-10
 Sequence 10. Application US/06295814E
 Patent No. 5658786
 GENERAL INFORMATION:
 APPLICANT: Smith, Kelli E.
 APPLICANT: Borden, Laurence A.
 APPLICANT: Hartley, Paul R.
 APPLICANT: Weinstein, Richard L.
 TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
 TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.30
 CURRENT APPLICATION DATA:

MOLECULE TYPE: protein
US-08-095-734-2

Query Match 26.5%; Score 44; DB 1; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 2 WRSQSSVFGQLFGQHPDVFYIMEPA 26
DB 116 WRTGSPTPGARGHPPEAIDLAEDA 140

RESULT 38

US-08-444-623-2
Sequence 2, Application US/08444623
Patent No. 5866403
GENERAL INFORMATION:
APPLICANT: Aldovini, Anna
APPLICANT: Young, Richard A.
TITLE OF INVENTION: Homologously Recombinant Slow Growing
TITLE OF INVENTION: Mycobacteria and Uses Therefor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,623
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/095,734
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: 07/711,334
FILING DATE: 06-JUN-1991
APPLICATION NUMBER: 07/367,894
FILING DATE: 19-JUN-1989
APPLICATION NUMBER: PCT/US90/03451
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: PCT/US89/02962
FILING DATE: 07-JUL-1989
APPLICATION NUMBER: 07/361,944
FILING DATE: 05-JUN-1989
APPLICATION NUMBER: 07/223,089
FILING DATE: 22-JULY-1988
APPLICATION NUMBER: 07/216,390
FILING DATE: 07-JUL-1988
APPLICATION NUMBER: 07/163,546
FILING DATE: 03-MAR-1988
APPLICATION NUMBER: PCT/US88/00614
FILING DATE: 29-FEB-1988
APPLICATION NUMBER: 07/020,451
FILING DATE: 02-MAR-1987
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH193-11M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1271 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-623-2

Query Match 26.5%; Score 44; DB 2; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 2 WRSQSSVFGQLFGQHPDVFYIMEPA 26
DB 116 WRTGSPTPGARGHPPEAIDLAEDA 140

RESULT 39

US-08-471-869-2
Sequence 2, Application US/08471869
Patent No. 6022745
GENERAL INFORMATION:
APPLICANT: Aldovini, Anna
APPLICANT: Young, Richard A.
TITLE OF INVENTION: Homologously Recombinant Slow Growing
TITLE OF INVENTION: Mycobacteria and Uses Therefor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,869
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/08267
FILING DATE: 22-JUL-1994
APPLICATION NUMBER: 08/095,734
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: 07/711,334
FILING DATE: 06-JUN-1991
APPLICATION NUMBER: US 07/711,334
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/367,894
FILING DATE: 19-JUN-1989
APPLICATION NUMBER: PCT/US90/03451
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: PCT/US89/02962
FILING DATE: 07-JUL-1989
APPLICATION NUMBER: US 07/361,944
FILING DATE: 05-JUN-1989
APPLICATION NUMBER: PCT/US88/00614
FILING DATE: 29-FEB-1988
APPLICATION NUMBER: 07/223,089
FILING DATE: 22-JUL-1988
APPLICATION NUMBER: US 07/216,390
FILING DATE: 07-JUL-1988
APPLICATION NUMBER: US 07/163,546
FILING DATE: 03-MAR-1988
APPLICATION NUMBER: PCT/US88/00614
FILING DATE: 29-FEB-1988
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/020,451
; FILING DATE: 02-MAR-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH193-11MA2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-869-2

Query Match      26.5%; Score 44; DB 3; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY      2 WRGSSFFVGLFGQHPDVFYLMPEA 26
Db      116 WRTGSPTPGPARGHPPEAIDLAEDA 140

RESULT 40
; US-09-342-563-2
; Sequence 2, Application US/09342563
; Patent No. 6355486
; GENERAL INFORMATION:
; APPLICANT: Aldovini, Anna
; APPLICANT: Young, Richard A.
; TITLE OF INVENTION: Homologously Recombinant Slow Growing
; TITLE OF INVENTION: Mycobacteria and Uses Therefor
; FILE REFERENCE: WH193-11MA3
; CURRENT APPLICATION NUMBER: US/09/342,563
; EARLIER FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/471,869
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: US 08/095,734
; EARLIER FILING DATE: 1993-07-22
; EARLIER APPLICATION NUMBER: PCT/US94/08267
; EARLIER FILING DATE: 1994-07-22
; EARLIER APPLICATION NUMBER: US 07/11,334
; EARLIER FILING DATE: 1991-06-06
; EARLIER APPLICATION NUMBER: US 07/367,894
; EARLIER FILING DATE: 1989-06-19
; EARLIER APPLICATION NUMBER: PCT/US90/03451
; EARLIER FILING DATE: 1990-06-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1271
; TYPE: PRT
; ORGANISM: Mycobacterium bovis Bacille Calmette-Guerin
; US-09-342-563-2

Query Match      26.5%; Score 44; DB 4; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY      2 WRGSSFFVGLFGQHPDVFYLMPEA 26
Db      116 WRTGSPTPGPARGHPPEAIDLAEDA 140

TITLE OF INVENTION: Homologously Recombinant Slow Growing
TITLE OF INVENTION: Mycobacteria and Uses Therefor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/08267
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/095,734
FILING DATE: 22-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH193-11MA PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1271 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-08267-2

Query Match      26.5%; Score 44; DB 5; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY      2 WRGSSFFVGLFGQHPDVFYLMPEA 26
Db      116 WRTGSPTPGPARGHPPEAIDLAEDA 140

RESULT 42
; US-08-469-005A-10
; Sequence 10, Application US/08469005A
; Patent No. 565874
; GENERAL INFORMATION:
; APPLICANT: KUHADA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/188,426
FILING DATE: 24-JAN-1994
APPLICATION NUMBER: 08/096,908
FILING DATE: 26-JUL-1993
APPLICATION NUMBER: 07/917,716
FILING DATE: 24-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H
REGISTRATION NUMBER: 34,638
REFERENCE/DOCKET NUMBER: 062482-0113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2509 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-469-005A-10

Query Match 26.5%; Score 44; DB 1; Length 2509;
Best Local Similarity 26.3%; Pred. No. 5.4e+02;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SFFVGLFGQHPDVFYIMEP 25
DB 40 WKAGLYGLPRSGKLKLDLSRFASFFGVHPKQAHMDP 77

RESULT 43
US-09-261-907-2
Sequence 2, Application US/09261907A
Patent No. 6294364
GENERAL INFORMATION:
APPLICANT: ELLIS, CATHERINE
APPLICANT: LONSDALE, JOHN
APPLICANT: BERGSMAN, DEBK J.
APPLICANT: MOONEY, JEFFREY L.
APPLICANT: DEPIERA, MEGAN E.
APPLICANT: CHAPMAN, CONRAD
TITLE OF INVENTION: HUMAN FAS
FILE REFERENCE: GP-70603
CURRENT APPLICATION NUMBER: US/09/261.907A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2511
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-261-907-2

Query Match 26.5%; Score 44; DB 4; Length 2511;
Best Local Similarity 26.3%; Pred. No. 5.4e+02;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SFFVGLFGQHPDVFYIMEP 25
DB 40 WKAGLYGLPRSGKLKLDLSRFASFFGVHPKQAHMDP 77

RESULT 44
US-08-906-769-89
Sequence 89, Application US/08906769
Patent No. 6077687
GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-769-89

Query Match 25.9%; Score 43; DB 3; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWRGSSFFVGLFGQHPDVFYIMEPAW 27
DB 75 SVRVGTSPQGRGVSHPXAQIKHPAY 101

RESULT 45
US-08-906-616-89
Sequence 89, Application US/08906616
Patent No. 6121035
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA

ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,616
FILING DATE: 05-AUG-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-616-89

Query Match 25.9%; Score 43; DB 3; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SRRSGSFVQGLFGQHPDVFYLMPEAW 27
DB 75 SVRVGTSFGGRGSRVHPXAQIIKHPAY 101

RESULT 46
US-08-817-795-89
Sequence 89, Application US/08817795
Patent No. 6139840
GENERAL INFORMATION:
APPLICANT: Grievie, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Hunter, Shirley Wu
APPLICANT: Frank, Glenn R.
APPLICANT: Health, Andrew W.
APPLICANT: Yamaka, Miles Yamataka
APPLICANT: Arfsten, Ann
APPLICANT: Dale, Beverly
APPLICANT: Stiegler, Gary
TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES, AND USES THEREOF
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,795
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14442
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Gary J. Connell
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-795-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SRRSGSFVQGLFGQHPDVFYLMPEAW 27
DB 75 SVRVGTSFGGRGSRVHPXAQIIKHPAY 101

RESULT 47
US-08-639-075A-89
Sequence 89, Application US/08639075A
Patent No. 6150125
GENERAL INFORMATION:
APPLICANT: Grievie, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-639-075A-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-613-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
QY 1 SWRGSFVQLFQGHDPDYELMPEAW 27
DB 75 SVRVGTSTFGGRGSHVHPXAGTIKHPAY 101

Search completed: February 20, 2003, 13:35:00
Job time : 49.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:33:52 ; Search time 13.2571 Seconds
(without alignments)
55.888 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166
Sequence: 1 SWRGSSEWGOLFQHPDVFLMEPMHV 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 segs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/PC1US_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppa/PC1US_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	166	100.0	171 10 US-09-827-602-8	Sequence 8, Appl1
2	166	100.0	386 9 US-10-007-262-1	Sequence 1, Appl1
3	166	100.0	386 10 US-09-816-825-2	Sequence 2, Appl1
4	160	96.4	169 10 US-09-927-602-6	Sequence 6, Appl1
5	160	96.4	395 10 US-09-927-602-2	Sequence 2, Appl1
6	160	96.4	395 10 US-09-927-602-3	Sequence 3, Appl1
7	160	96.4	418 10 US-09-927-602-5	Sequence 5, Appl1
8	157	94.6	169 10 US-09-827-602-7	Sequence 7, Appl1
9	157	94.6	390 10 US-09-927-602-4	Sequence 4, Appl1
10	124	74.7	169 10 US-09-927-602-9	Sequence 9, Appl1
11	124	74.7	483 9 US-10-212-933-2	Sequence 2, Appl1
12	124	74.7	484 9 US-09-833-790-255	Sequence 255, Appl1
13	124	74.7	531 10 US-09-927-602-10	Sequence 10, Appl1
14	116	69.9	174 10 US-09-927-602-11	Sequence 11, Appl1
15	101	60.8	344 9 US-09-738-626-5973	Sequence 5973, Ap
16	50	30.1	265 10 US-09-888-623-2	Sequence 2, Appl1
17	49	29.5	265 10 US-09-888-623-14	Sequence 14, Appl1
18	49	29.5	318 10 US-09-888-623-16	Sequence 16, Appl1
19	49	29.5		

20	48	28.9	460 10 US-09-815-242-5057	Sequence 5057, Ap
21	48	28.9	753 10 US-09-888-615-68	Sequence 68, Appl1
22	47	28.3	701 10 US-09-815-242-12050	Sequence 12050, A
23	47	28.3	925 9 US-09-738-626-3608	Sequence 3608, Ap
24	46	27.7	302 9 US-09-738-626-3746	Sequence 3746, Ap
25	45.5	27.4	844 10 US-09-272-809-6	Sequence 6, Appl1
26	45	27.1	99 10 US-09-864-761-41772	Sequence 41772, A
27	45	27.1	239 9 US-09-925-299-689	Sequence 834, Ap
28	45	27.1	376 9 US-09-738-626-6189	Sequence 6189, Ap
29	45	27.1	791 10 US-09-735-705-170	Sequence 170, Ap
30	45	27.1	791 10 US-09-850-716A-170	Sequence 170, Ap
31	45	27.1	791 10 US-09-897-778-170	Sequence 170, Ap
32	45	27.1	847 9 US-10-112-527-4	Sequence 4, Appl1
33	45	27.1	881 10 US-09-850-716A-430	Sequence 430, Ap
34	45	27.1	881 10 US-09-735-705-161	Sequence 430, Ap
35	45	27.1	920 10 US-09-735-705-357	Sequence 357, Ap
36	45	27.1	920 10 US-09-850-716A-357	Sequence 357, Ap
37	45	27.1	920 10 US-09-897-778-357	Sequence 357, Ap
38	45	27.1	942 9 US-09-924-298-71	Sequence 71, Appl1
39	45	27.1	942 10 US-09-919-172-87	Sequence 87, Appl1
40	45	27.1	943 10 US-09-735-705-161	Sequence 161, Ap
41	45	27.1	943 10 US-09-850-716A-161	Sequence 161, Ap
42	45	27.1	943 10 US-09-897-778-161	Sequence 161, Ap
43	44	26.5	812 10 US-09-815-242-11294	Sequence 11294, A
44	43.5	26.2	1054 9 US-09-738-626-5852	Sequence 5852, Ap
45	43	25.9	466 10 US-09-925-302-729	Sequence 729, Ap
46	43	25.9	1160 10 US-09-815-242-5480	Sequence 5480, Ap
47	43	25.9	1168 10 US-09-815-242-1151	Sequence 1151, A
48	42.5	25.6	80 9 US-09-764-870-510	Sequence 510, Ap
49	42.5	25.6	94 10 US-10-001-835-184	Sequence 184, Ap
50	42.5	25.6	229 9 US-09-922-598-281	Sequence 281, Ap
51	42.5	25.6	229 9 US-09-989-735-281	Sequence 281, Ap
52	42.5	25.6	229 9 US-09-989-735-281	Sequence 281, Ap
53	42.5	25.6	229 9 US-09-989-735-281	Sequence 281, Ap
54	42.5	25.6	229 9 US-09-989-735-281	Sequence 281, Ap
55	42.5	25.6	229 9 US-09-990-436-281	Sequence 281, Ap
56	42.5	25.6	229 9 US-09-991-181-281	Sequence 281, Ap
57	42.5	25.6	229 9 US-09-993-687-281	Sequence 281, Ap
58	42.5	25.6	229 9 US-09-989-734-281	Sequence 281, Ap
59	42.5	25.6	229 9 US-09-997-653-281	Sequence 281, Ap
60	42.5	25.6	229 9 US-09-993-667-281	Sequence 281, Ap
61	42.5	25.6	229 9 US-09-990-438-281	Sequence 281, Ap
62	42.5	25.6	229 9 US-09-990-562-281	Sequence 281, Ap
63	42.5	25.6	229 9 US-09-997-428-281	Sequence 281, Ap
64	42.5	25.6	229 9 US-09-997-666-281	Sequence 281, Ap
65	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
66	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
67	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
68	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
69	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
70	42.5	25.6	229 10 US-09-989-722-281	Sequence 281, Ap
71	42.5	25.6	229 10 US-09-991-073-281	Sequence 281, Ap
72	42.5	25.6	229 10 US-09-990-442-281	Sequence 281, Ap
73	42.5	25.6	229 10 US-09-991-163-281	Sequence 281, Ap
74	42.5	25.6	229 10 US-09-993-604-281	Sequence 281, Ap
75	42.5	25.6	229 10 US-09-990-456-281	Sequence 281, Ap
76	42.5	25.6	229 10 US-09-989-721-281	Sequence 281, Ap
77	42	25.3	66 10 US-09-867-550-1170	Sequence 1170, Ap
78	42	25.3	74 9 US-09-989-920-255	Sequence 255, Appl1
79	42	25.3	252 10 US-09-347-064-2	Sequence 2, Appl1
80	42	25.3	252 10 US-09-347-064-8	Sequence 8, Appl1
81	42	25.3	335 9 US-10-127-884-170	Sequence 170, Ap
82	42	25.3	335 9 US-09-924-400-304	Sequence 304, Ap
83	42	25.3	334 9 US-09-924-400-333	Sequence 333, Ap
84	42	25.3	334 9 US-09-924-400-340	Sequence 340, Ap
85	42	25.3	384 10 US-09-825-301-8	Sequence 8, Appl1
86	42	25.3	384 10 US-09-810-936-334	Sequence 304, Ap
87	42	25.3	384 10 US-09-810-936-333	Sequence 333, Ap
88	42	25.3	384 10 US-09-429-755-364	Sequence 364, Ap
89	42	25.3	334 9 US-09-924-400-336	Sequence 336, Ap
90	42	25.3	559 9 US-09-924-400-324	Sequence 324, Ap
91	42	25.3	559 10 US-09-810-936-324	Sequence 324, Ap
92	42	25.3	666 9 US-09-924-400-305	Sequence 305, Ap

93	42	25.3	656	9	US-10-012-896-379	Sequence 379, App	156	40	24.1	1491	10	US-09-815-242-5568	Sequence 5568, App
94	42	25.3	656	9	US-09-895-793-379	Sequence 379, App	157	40	24.1	1502	10	US-09-815-242-12162	Sequence 12162, A
95	42	25.3	656	9	US-09-895-814-379	Sequence 379, App	158	39.5	23.8	1248	10	US-09-825-999-958	Sequence 958, App
96	42	25.3	656	10	US-09-825-301-9	Sequence 9, Appli	159	39.5	23.8	323	10	US-09-903-395-2	Sequence 2, Appli
97	42	25.3	656	10	US-09-758-143-379	Sequence 379, App	160	39.5	23.8	333	10	US-09-828-113-33	Sequence 33, Appli
98	42	25.3	656	10	US-09-780-669-379	Sequence 379, App	161	39.5	23.8	337	9	US-09-860-846-14	Sequence 14, Appli
99	42	25.3	656	10	US-09-810-936-305	Sequence 305, App	162	39.5	23.8	337	10	US-09-861-889-14	Sequence 14, Appli
100	42	25.3	656	10	US-09-828-827-379	Sequence 379, App	163	39.5	23.8	500	9	US-09-712-363-774	Sequence 274, App
101	42	25.3	656	10	US-09-429-755-305	Sequence 305, App	164	39.5	23.8	877	9	US-10-043-418-1	Sequence 1, Appli
102	42	25.3	671	9	US-09-924-400-306	Sequence 306, App	165	39.5	23.8	1169	9	US-09-815-242-13448	Sequence 13448, A
103	42	25.3	671	9	US-10-012-896-380	Sequence 380, App	166	39.5	23.8	1169	10	US-09-815-242-13693	Sequence 13693, A
104	42	25.3	671	9	US-09-895-793-380	Sequence 380, App	167	39.5	23.8	3782	9	US-09-815-242-13693	Sequence 4, Appli
105	42	25.3	671	9	US-09-895-814-380	Sequence 380, App	168	39.5	23.8	3782	10	US-09-860-846-4	Sequence 4, Appli
106	42	25.3	671	10	US-09-825-301-10	Sequence 10, Appli	169	39.5	23.5	17	9	US-09-996-634-91	Sequence 9, Appli
107	42	25.3	671	10	US-09-759-143-380	Sequence 380, App	170	39.5	23.5	63	10	US-09-864-761-42380	Sequence 42380, A
108	42	25.3	671	10	US-09-780-669-380	Sequence 380, App	181	39.5	23.5	111	10	US-09-925-300-980	Sequence 980, App
109	42	25.3	671	10	US-09-810-936-306	Sequence 306, App	182	39.5	23.5	154	10	US-09-764-864-1156	Sequence 1156, App
110	42	25.3	671	10	US-09-828-827-380	Sequence 380, App	183	39.5	23.5	154	10	US-09-764-864-1569	Sequence 1569, App
111	42	25.3	671	10	US-09-429-755-306	Sequence 306, App	184	39.5	23.5	155	10	US-09-729-674-104	Sequence 104, App
112	42	25.3	876	10	US-09-815-242-12623	Sequence 12623, A	185	39.5	23.5	171	9	US-09-895-913A-26	Sequence 26, Appli
113	42	25.3	1719	9	US-10-012-896-378	Sequence 378, App	186	39.5	23.5	178	10	US-09-864-761-41969	Sequence 41969, A
114	42	25.3	1719	9	US-09-895-793-378	Sequence 378, App	187	39.5	23.5	284	10	US-09-771-161A-106	Sequence 106, App
115	42	25.3	1719	9	US-09-895-814-378	Sequence 378, App	188	39.5	23.5	299	10	US-09-886-055-41	Sequence 41, Appli
116	42	25.3	1719	10	US-09-759-143-378	Sequence 378, App	189	39.5	23.5	300	10	US-09-919-770-2	Sequence 2, Appli
117	42	25.3	1719	10	US-09-780-669-378	Sequence 378, App	190	39.5	23.5	305	10	US-09-925-300-51205	Sequence 1205, App
118	42	25.3	1719	10	US-09-828-827-378	Sequence 378, App	191	39.5	23.5	314	10	US-09-925-300-51205	Sequence 2, Appli
119	41.5	25.0	56	9	US-09-796-692-748	Sequence 748, App	192	39.5	23.5	318	10	US-09-950-368-16	Sequence 16, Appli
120	41.5	25.0	105	10	US-09-864-761-33632	Sequence 33632, A	193	39.5	23.5	322	10	US-09-942-447-5	Sequence 5, Appli
121	41.5	25.0	210	12	US-10-029-079-3	Sequence 3, Appli	194	39.5	23.5	365	10	US-09-818-512-4	Sequence 4, Appli
122	41.5	25.0	730	9	US-09-978-295A-102	Sequence 102, App	195	39.5	23.5	370	10	US-10-114-693-174	Sequence 2, Appli
123	41.5	25.0	730	9	US-09-978-697-102	Sequence 102, App	196	39.5	23.5	370	10	US-09-849-031A-2	Sequence 2, Appli
124	41.5	25.0	730	9	US-09-978-192A-102	Sequence 102, App	197	39.5	23.5	370	10	US-09-849-562A-2	Sequence 2, Appli
125	41.5	25.0	730	9	US-09-999-832A-102	Sequence 102, App	198	39.5	23.5	376	10	US-09-818-512-2	Sequence 2, Appli
126	41.5	25.0	730	9	US-09-978-188-102	Sequence 102, App	199	39.5	23.5	444	10	US-09-881-752A-212	Sequence 212, App
127	41.5	25.0	349	10	US-09-916-031-1	Sequence 1, Appli	200	39.5	23.5	477	9	US-09-738-626-7048	Sequence 5048, App
128	41.5	25.0	378	9	US-09-916-031-1	Sequence 1, Appli	201	39.5	23.5	503	9	US-09-823-038A-51	Sequence 51, Appli
129	41.5	25.0	378	9	US-09-759-130B-76	Sequence 76, Appli	202	39.5	23.5	503	10	US-09-823-038A-51	Sequence 2, Appli
130	41.5	25.0	429	9	US-09-922-683-9	Sequence 9, Appli	203	39.5	23.5	547	10	US-09-895-860-2	Sequence 2, Appli
131	41.5	25.0	429	9	US-09-922-683-9	Sequence 9, Appli	204	39.5	23.5	551	10	US-09-802-472A-21	Sequence 21, Appli
132	41.5	25.0	429	9	US-09-922-683-9	Sequence 9, Appli	205	39.5	23.5	571	9	US-10-036-041-23	Sequence 23, Appli
133	41.5	25.0	511	10	US-09-925-301-1060	Sequence 1060, App	206	39.5	23.5	571	9	US-10-028-072-542	Sequence 542, App
134	41.5	25.0	530	10	US-09-815-242-10105	Sequence 10105, A	207	39.5	23.5	571	9	US-10-035-855-23	Sequence 23, Appli
135	41.5	25.0	531	10	US-09-815-242-10105	Sequence 10105, A	208	39.5	23.5	571	9	US-10-121-049-542	Sequence 542, App
136	41.5	25.0	540	10	US-09-729-454-32	Sequence 32, Appli	209	39.5	23.5	571	9	US-10-123-004-542	Sequence 542, App
137	41.5	25.0	681	9	US-09-759-130B-75	Sequence 75, Appli	210	39.5	23.5	571	9	US-10-140-470-542	Sequence 542, App
138	41.5	25.0	715	9	US-10-003-132-2	Sequence 2, Appli	211	39.5	23.5	571	9	US-09-931-836-23	Sequence 23, Appli
139	41.5	25.0	715	9	US-09-759-130B-73	Sequence 73, Appli	212	39.5	23.5	571	9	US-10-175-746-542	Sequence 542, App
140	41.5	25.0	906	10	US-09-815-242-10567	Sequence 10567, A	213	39.5	23.5	571	9	US-10-176-518-542	Sequence 542, App
141	40.5	24.4	624	9	US-09-738-626-4289	Sequence 4289, App	214	39.5	23.5	571	9	US-10-176-518-542	Sequence 542, App
142	40.5	24.4	5215	9	US-09-860-846-2	Sequence 2, Appli	215	39.5	23.5	571	9	US-10-176-518-542	Sequence 542, App
143	40.5	24.4	5215	10	US-09-861-289-2	Sequence 2, Appli	216	39.5	23.5	571	12	US-10-036-342-23	Sequence 23, Appli
144	40.5	24.4	17	9	US-09-996-634-90	Sequence 90, Appli	217	39.5	23.5	572	10	US-09-828-302-11	Sequence 11, Appli
145	40.5	24.4	246	9	US-09-909-567B-49	Sequence 49, Appli	218	39.5	23.5	711	10	US-09-765-272-82	Sequence 82, Appli
146	40.5	24.4	249	9	US-10-136-517-17	Sequence 17, Appli	219	39.5	23.5	777	9	US-10-174-590-348	Sequence 348, App
147	40.5	24.4	263	10	US-09-888-623-3	Sequence 3, Appli	220	39.5	23.5	777	9	US-10-174-590-348	Sequence 348, App
148	40.5	24.4	377	9	US-09-738-626-3873	Sequence 3873, App	221	39.5	23.5	777	9	US-10-176-758-348	Sequence 348, App
149	40.5	24.4	403	9	US-10-034-158-3	Sequence 3, Appli	222	39.5	23.5	777	9	US-10-175-737-348	Sequence 348, App
150	40.5	24.4	403	10	US-09-086-118-23	Sequence 23, Appli	223	39.5	23.5	777	9	US-10-173-706-348	Sequence 348, App
151	40.5	24.4	403	10	US-09-835-684-3	Sequence 3, Appli	224	39.5	23.5	777	9	US-10-175-738-348	Sequence 348, App
152	40.5	24.4	403	10	US-09-880-371-3	Sequence 3, Appli	225	39.5	23.5	777	9	US-10-175-738-348	Sequence 348, App
153	40.5	24.4	403	10	US-09-879-248-3	Sequence 3, Appli	226	39.5	23.5	777	9	US-10-176-482-348	Sequence 348, App
154	40.5	24.4	403	10	US-09-770-693-3	Sequence 3, Appli	227	39.5	23.5	777	9	US-10-176-482-348	Sequence 348, App
155	40.5	24.4	403	10	US-09-766-348-3	Sequence 3, Appli	228	39.5	23.5	777	9	US-10-176-482-348	Sequence 348, App
156	40.5	24.4	423	10	US-09-733-009-142	Sequence 142, App	229	39.5	23.5	777	9	US-10-180-552-348	Sequence 348, App
157	40.5	24.4	423	10	US-09-793-306-142	Sequence 142, App	230	39.5	23.5	777	9	US-10-173-700-348	Sequence 348, App
158	40.5	24.4	488	9	US-09-996-634-140	Sequence 140, App	231	39.5	23.5	777	9	US-10-174-579-348	Sequence 348, App
159	40.5	24.4	513	9	US-09-738-626-4772	Sequence 4772, App	232	39.5	23.5	777	9	US-10-174-579-348	Sequence 348, App
160	40.5	24.4	548	9	US-09-738-626-3855	Sequence 3855, App	233	39.5	23.5	777	9	US-10-174-582-348	Sequence 348, App
161	40.5	24.4	710	10	US-09-287-849-16	Sequence 16, Appli	234	39.5	23.5	777	9	US-10-174-588-348	Sequence 348, App
162	40.5	24.4	788	10	US-09-391-340-6	Sequence 6, Appli	235	39.5	23.5	777	9	US-10-175-739-348	Sequence 348, App
163	40.5	24.4	788	10	US-09-948-369-6	Sequence 6, Appli	236	39.5	23.5	777	9	US-10-175-740-348	Sequence 348, App
164	40.5	24.4	856	10	US-09-287-849-12	Sequence 12, Appli	237	39.5	23.5	777	9	US-10-175-743-348	Sequence 348, App
165	40.5	24.4	940	9	US-09-811-088-11	Sequence 11, Appli	238	39.5	23.5	777	9	US-10-176-488-348	Sequence 348, App

229	39	23.5	777	9	US-10-176-492-348	Sequence 348, App	312	38	22.9	467	9	US-10-176-513-16	Sequence 16, Appl
240	39	23.5	777	9	US-10-176-747-348	Sequence 348, App	313	38	22.9	467	9	US-10-180-552-16	Sequence 16, Appl
241	39	23.5	777	9	US-10-176-750-348	Sequence 348, App	314	38	22.9	467	9	US-10-180-557-16	Sequence 16, Appl
242	39	23.5	777	9	US-10-176-985-348	Sequence 348, App	315	38	22.9	467	9	US-09-148-545-134	Sequence 134, App
243	39	23.5	777	9	US-10-176-987-348	Sequence 348, App	316	38	22.9	467	9	US-09-148-638-195	Sequence 195, App
244	39	23.5	777	9	US-10-176-991-348	Sequence 348, App	317	38	22.9	467	9	US-09-907-613-195	Sequence 195, App
245	39	23.5	777	9	US-10-176-992-348	Sequence 348, App	318	38	22.9	467	9	US-09-907-642-195	Sequence 195, App
246	39	23.5	777	9	US-10-176-993-348	Sequence 348, App	319	38	22.9	467	9	US-10-173-700-16	Sequence 16, Appl
247	39	23.5	777	9	US-10-184-658-348	Sequence 348, App	320	38	22.9	467	9	US-10-174-572-16	Sequence 16, Appl
248	39	23.5	777	12	US-10-052-856-348	Sequence 348, App	321	38	22.9	467	9	US-10-174-579-16	Sequence 16, Appl
249	39	23.5	832	10	US-09-729-674-176	Sequence 176, App	322	38	22.9	467	9	US-10-174-582-16	Sequence 16, Appl
250	39	23.5	966	10	US-09-771-1616-157	Sequence 197, App	323	38	22.9	467	9	US-10-174-588-16	Sequence 16, Appl
251	38.5	23.2	102	10	US-09-864-761-37086	Sequence 37086, A	324	38	22.9	467	9	US-10-175-730-16	Sequence 16, Appl
252	38.5	23.2	233	9	US-09-975-719-81	Sequence 18, Appl	325	38	22.9	467	9	US-10-175-749-16	Sequence 16, Appl
253	38.5	23.2	273	9	US-10-063-547-168	Sequence 18, Appl	326	38	22.9	467	9	US-10-175-743-16	Sequence 16, Appl
254	38.5	23.2	273	9	US-10-036-041-18	Sequence 50, App	327	38	22.9	467	9	US-10-176-488-16	Sequence 16, Appl
255	38.5	23.2	273	9	US-10-028-072-640	Sequence 18, Appl	328	38	22.9	467	9	US-10-176-492-16	Sequence 16, Appl
256	38.5	23.2	273	9	US-10-035-855-18	Sequence 18, Appl	329	38	22.9	467	9	US-10-176-747-16	Sequence 16, Appl
257	38.5	23.2	273	9	US-10-063-616-168	Sequence 18, App	330	38	22.9	467	9	US-10-176-750-16	Sequence 16, Appl
258	38.5	23.2	273	9	US-10-063-502-168	Sequence 168, App	331	38	22.9	467	9	US-10-176-985-16	Sequence 16, Appl
259	38.5	23.2	273	9	US-10-121-049-540	Sequence 540, App	332	38	22.9	467	9	US-10-176-987-16	Sequence 16, Appl
260	38.5	23.2	273	9	US-10-123-904-540	Sequence 540, App	333	38	22.9	467	9	US-10-176-991-16	Sequence 16, Appl
261	38.5	23.2	273	9	US-10-140-470-540	Sequence 540, App	334	38	22.9	467	9	US-10-176-992-16	Sequence 16, Appl
262	38.5	23.2	273	9	US-09-931-836-18	Sequence 18, Appl	335	38	22.9	467	9	US-10-176-993-16	Sequence 16, Appl
263	38.5	23.2	273	9	US-10-175-746-40	Sequence 540, App	336	38	22.9	467	9	US-10-184-938-16	Sequence 16, Appl
264	38.5	23.2	273	9	US-10-176-918-540	Sequence 540, App	337	38	22.9	467	9	US-09-909-320-195	Sequence 195, App
265	38.5	23.2	273	9	US-10-176-921-540	Sequence 540, App	338	38	22.9	467	10	US-09-909-320-195	Sequence 195, App
266	38.5	23.2	273	12	US-10-036-342-18	Sequence 18, App	339	38	22.9	467	12	US-10-052-586-16	Sequence 16, Appl
267	38.5	23.2	273	12	US-10-006-667-168	Sequence 168, App	340	38	22.9	467	9	US-09-965-529-16	Sequence 16, Appl
268	38.5	23.2	465	9	US-10-029-180-44	Sequence 44, App	341	38	22.9	467	9	US-10-016-283-36	Sequence 12, Appl
269	38.5	23.2	1150	10	US-09-864-739-9	Sequence 9, Appl	342	38	22.9	467	10	US-09-843-845-2	Sequence 12, Appl
270	38	22.9	26	10	US-09-864-761-41030	Sequence 41030, A	343	38	22.9	520	10	US-09-765-272-122	Sequence 12, Appl
271	38	22.9	66	10	US-09-872-523-66	Sequence 66, Appl	344	38	22.9	521	10	US-09-815-242-11855	Sequence 11855, A
272	38	22.9	66	10	US-09-925-300-1832	Sequence 1832, App	345	38	22.9	534	9	US-09-738-626-5899	Sequence 5899, App
273	38	22.9	71	10	US-09-894-882-74	Sequence 74, Appl	346	38	22.9	539	10	US-09-945-826-10	Sequence 10, Appl
274	38	22.9	81	9	US-09-861-097-25	Sequence 25, Appl	347	38	22.9	542	10	US-09-805-456A-4	Sequence 4, Appl
275	38	22.9	110	10	US-09-864-761-43917	Sequence 43917, A	348	38	22.9	549	10	US-09-805-456A-2	Sequence 2, Appl
276	38	22.9	162	9	US-09-796-662-2342	Sequence 2342, App	349	38	22.9	658	10	US-09-815-242-11824	Sequence 11824, A
277	38	22.9	177	9	US-10-105-427-2	Sequence 2, Appl	350	38	22.9	788	9	US-10-029-180-92	Sequence 92, Appl
278	38	22.9	215	10	US-09-912-020-362	Sequence 362, App	351	38	22.9	803	10	US-09-968-436B-2	Sequence 2, Appl
279	38	22.9	216	9	US-10-016-283-31	Sequence 31, Appl	352	38	22.9	833	10	US-09-759-010-7	Sequence 7, Appl
280	38	22.9	256	9	US-10-016-283-30	Sequence 30, Appl	353	38	22.9	848	9	US-09-875-353-5	Sequence 5, Appl
281	38	22.9	294	9	US-10-016-283-29	Sequence 29, Appl	354	38	22.9	963	10	US-09-886-466-22	Sequence 22, Appl
282	38	22.9	316	10	US-09-815-242-13929	Sequence 13929, A	355	38	22.9	1037	10	US-09-728-727-55	Sequence 55, Appl
283	38	22.9	338	9	US-10-016-283-28	Sequence 28, Appl	356	38	22.9	1040	10	US-09-765-272-118	Sequence 118, App
284	38	22.9	338	9	US-10-034-158-1	Sequence 1, Appl	357	38	22.9	1157	10	US-09-815-242-11522	Sequence 11522, A
285	38	22.9	338	10	US-09-086-118-21	Sequence 21, Appl	358	38	22.9	1159	9	US-09-738-626-4140	Sequence 4140, App
286	38	22.9	338	10	US-09-835-684-1	Sequence 1, Appl	359	38	22.9	2905	8	US-08-424-950B-401	Sequence 401, App
287	38	22.9	338	10	US-09-880-771-1	Sequence 1, Appl	360	37.5	22.6	50	10	US-09-764-877-1816	Sequence 1816, App
288	38	22.9	338	10	US-09-879-248-1	Sequence 1, Appl	361	37.5	22.6	76	10	US-09-864-761-36015	Sequence 36015, A
289	38	22.9	338	10	US-09-770-693-1	Sequence 1, Appl	362	37.5	22.6	105	10	US-09-925-299-927	Sequence 927, App
290	38	22.9	338	10	US-09-766-348-1	Sequence 1, Appl	363	37.5	22.6	139	10	US-09-815-242-11799	Sequence 11799, A
291	38	22.9	362	9	US-09-796-753-142	Sequence 142, App	364	37.5	22.6	226	10	US-09-741-668-899	Sequence 346, App
292	38	22.9	390	9	US-10-016-283-27	Sequence 27, Appl	365	37.5	22.6	475	9	US-09-764-868-899	Sequence 899, App
293	38	22.9	419	10	US-09-740-033-4	Sequence 4, Appl	366	37.5	22.6	640	10	US-09-815-242-11884	Sequence 11884, A
294	38	22.9	440	9	US-10-016-283-26	Sequence 26, Appl	367	37.5	22.6	848	10	US-09-839-185-8	Sequence 8, Appl
295	38	22.9	456	9	US-10-016-283-25	Sequence 25, Appl	368	37.5	22.6	1337	10	US-09-803-126-1	Sequence 1, Appl
296	38	22.9	463	9	US-09-738-626-6946	Sequence 6946, App	369	37.5	22.6	19	10	US-09-864-761-45551	Sequence 45551, A
297	38	22.9	467	9	US-09-905-291A-195	Sequence 195, App	370	37	22.3	33	10	US-09-864-761-45594	Sequence 45594, A
298	38	22.9	467	9	US-09-981-876-134	Sequence 134, App	371	37	22.3	34	10	US-09-864-761-35790	Sequence 35790, A
299	38	22.9	467	9	US-09-902-853-195	Sequence 195, App	372	37	22.3	120	10	US-09-917-265-21	Sequence 21, Appl
300	38	22.9	467	9	US-09-907-824-195	Sequence 195, App	373	37	22.3	129	9	US-09-726-839-10	Sequence 10, Appl
301	38	22.9	467	9	US-09-907-841-195	Sequence 195, App	374	37	22.3	132	9	US-10-092-243A-17	Sequence 17, Appl
302	38	22.9	467	9	US-09-904-011-195	Sequence 195, App	375	37	22.3	155	9	US-10-108-605-361	Sequence 361, App
303	38	22.9	467	9	US-10-174-590-16	Sequence 16, Appl	376	37	22.3	112	10	US-09-815-242-5069	Sequence 5069, App
304	38	22.9	467	9	US-10-176-758-16	Sequence 16, Appl	377	37	22.3	227	9	US-09-738-626-6345	Sequence 6345, App
305	38	22.9	467	9	US-10-175-737-16	Sequence 16, Appl	378	37	22.3	302	10	US-09-864-761-35790	Sequence 35790, A
306	38	22.9	467	9	US-09-906-742-195	Sequence 195, App	379	37	22.3	315	10	US-09-886-056-307	Sequence 307, App
307	38	22.9	467	9	US-10-173-706-16	Sequence 16, Appl	380	37	22.3	340	9	US-10-155-947-2	Sequence 2, Appl
308	38	22.9	467	9	US-10-173-706-16	Sequence 16, Appl	381	37	22.3	340	9	US-10-155-947-5	Sequence 5, Appl
309	38	22.9	467	9	US-10-175-752-16	Sequence 16, Appl	382	37	22.3	340	10	US-10-155-947-6	Sequence 6, Appl
310	38	22.9	467	9	US-10-176-482-16	Sequence 16, Appl	383	37	22.3	340	10	US-09-775-195-2	Sequence 2, Appl
311	38	22.9	467	9	US-10-176-757-16	Sequence 16, Appl	384	37	22.3	340	10	US-09-775-195-5	Sequence 5, Appl

385	37	22.3	340	10	US-09-775-195-6	Sequence 6, Appl1	458	36.5	22.0	832	9	US-10-033-297-4	Sequence 4, Appl1
386	37	22.3	342	10	US-09-815-242-11043	Sequence 11043, A	459	36.5	22.0	832	9	US-09-972-834-2	Sequence 2, Appl1
387	37	22.3	345	9	US-10-155-947-4	Sequence 4, Appl1	460	36.5	22.0	832	9	US-10-071-505-1	Sequence 1, Appl1
388	37	22.3	346	10	US-09-775-195-4	Sequence 2, Appl1	461	36.5	22.0	832	9	US-10-081-806-4	Sequence 4, Appl1
389	37	22.3	345	10	US-09-829-482-2	Sequence 17, Appl1	462	36.5	22.0	832	9	US-10-074-328-4	Sequence 8, Appl1
390	37	22.3	366	10	US-09-901-252-17	Sequence 1020, Ap	463	36.5	22.0	833	9	US-10-033-297-8	Sequence 6, Appl1
391	37	22.3	374	9	US-09-764-868-1020	Sequence 50, Appl	464	36.5	22.0	833	9	US-10-033-297-6	Sequence 6, Appl1
392	37	22.3	379	10	US-09-934-868-50	Sequence 4635, Ap	465	36.5	22.0	833	9	US-10-033-297-9	Sequence 7, Appl1
393	37	22.3	392	9	US-09-738-626-4635	Sequence 15, Appl	466	36.5	22.0	833	9	US-10-033-297-11	Sequence 8, Appl1
394	37	22.3	410	10	US-09-917-265-15	Sequence 24, Appl	467	36.5	22.0	833	9	US-10-081-806-8	Sequence 8, Appl1
395	37	22.3	410	10	US-09-917-265-24	Sequence 6944, Ap	468	36.5	22.0	833	9	US-10-074-328-8	Sequence 9, Appl1
396	37	22.3	426	9	US-09-738-626-6944	Sequence 251, App	469	36.5	22.0	1016	12	US-10-007-693-95	Sequence 95, Appl1
397	37	22.3	438	9	US-09-712-363-251	Sequence 3703, Ap	470	36	21.7	24	10	US-09-864-761-47163	Sequence 47163, A
398	37	22.3	467	10	US-09-863-547B-1	Sequence 313, App	471	36	21.7	35	10	US-09-864-882-45369	Sequence 45369, A
399	37	22.3	491	10	US-09-738-626-3703	Sequence 10504, A	472	36	21.7	48	10	US-09-864-761-45369	Sequence 18, Appl
400	37	22.3	532	10	US-09-205-658-313	Sequence 10504, A	473	36	21.7	63	10	US-09-925-442-18	Sequence 48204, A
401	37	22.3	648	10	US-09-815-242-10504	Sequence 13812, A	474	36	21.7	68	10	US-09-864-761-48204	Sequence 1067, Ap
402	37	22.3	654	10	US-09-815-242-13812	Sequence 5453, Ap	475	36	21.7	82	10	US-09-764-867-11067	Sequence 11, Appl
403	37	22.3	722	9	US-09-738-626-5453	Sequence 18, Appl	476	36	21.7	96	10	US-09-833-067-11	Sequence 783, App
404	37	22.3	732	9	US-09-921-496-18	Sequence 5, Appl	477	36	21.7	102	10	US-09-925-299-783	Sequence 158, App
405	37	22.3	732	10	US-09-739-010-5	Sequence 18, Appl	478	36	21.7	125	12	US-10-078-829-158	Sequence 31, Appl
406	37	22.3	732	10	US-09-874-923-18	Sequence 4131, Ap	479	36	21.7	149	10	US-09-939-980-484	Sequence 377, App
407	37	22.3	799	9	US-09-738-626-4131	Sequence 113, App	480	36	21.7	152	10	US-09-764-870-377	Sequence 342, App
408	37	22.3	802	10	US-09-758-386-3	Sequence 5418, Ap	481	36	21.7	158	10	US-09-764-860-342	Sequence 30, Appl
409	37	22.3	805	9	US-10-108-605-113	Sequence 5118, Ap	482	36	21.7	178	9	US-10-076-785-30	Sequence 2, Appl1
410	37	22.3	1100	10	US-09-815-242-11861	Sequence 6, Appl1	483	36	21.7	195	10	US-09-833-067-2	Sequence 4, Appl1
411	37	22.3	1283	9	US-09-738-626-5418	Sequence 3, Appl1	484	36	21.7	195	10	US-09-833-067-6	Sequence 6, Appl1
412	37	22.3	1317	10	US-09-815-242-5118	Sequence 3, Appl1	485	36	21.7	195	10	US-09-833-067-8	Sequence 32, Appl1
413	37	22.3	1430	10	US-09-740-274-6	Sequence 14, Appl	486	36	21.7	195	10	US-09-833-079-32	Sequence 36, Appl
414	37	22.3	2037	9	US-09-951-402-3	Sequence 33, Appl	487	36	21.7	195	10	US-09-833-079-36	Sequence 38, Appl
415	37	22.3	2037	10	US-09-922-101-3	Sequence 6, Appl1	488	36	21.7	195	10	US-09-833-079-39	Sequence 39, Appl
416	37	22.3	2037	9	US-09-922-101-3	Sequence 3, Appl1	489	36	21.7	195	10	US-09-978-697-477	Sequence 477, App
417	37	22.3	2224	10	US-10-115-563-14	Sequence 37259, A	490	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
418	37	22.3	3739	9	US-09-860-846-33	Sequence 1608, Ap	491	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
419	37	22.3	3739	10	US-09-861-289-33	Sequence 1417, Ap	492	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
420	37	22.3	3739	10	US-09-861-289-33	Sequence 1885, Ap	493	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
421	37	22.3	11877	10	US-09-861-289-6	Sequence 669, App	494	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
422	37	22.3	11877	10	US-09-861-289-6	Sequence 4931, Ap	495	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
423	36.5	22.0	65	10	US-09-864-761-37259	Sequence 1175, Ap	496	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
424	36.5	22.0	62	10	US-09-925-300-1608	Sequence 620, App	497	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
425	36.5	22.0	77	9	US-09-796-692-1417	Sequence 61, Appl	498	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
426	36.5	22.0	77	9	US-09-796-692-1417	Sequence 2, Appl1	499	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
427	36.5	22.0	79	9	US-09-796-692-1417	Sequence 1, Appl1	500	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
428	36.5	22.0	101	9	US-09-796-692-1417	Sequence 3, Appl1	501	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
429	36.5	22.0	105	10	US-09-764-869-1175	Sequence 268, App	502	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
430	36.5	22.0	240	9	US-09-989-920-247	Sequence 13180, A	503	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
431	36.5	22.0	305	10	US-09-925-302-620	Sequence 10200, A	504	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
432	36.5	22.0	320	9	US-10-033-297-61	Sequence 13821, A	505	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
433	36.5	22.0	339	10	US-09-943-671-2	Sequence 6778, Ap	506	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
434	36.5	22.0	339	12	US-10-067-291-1	Sequence 16, Appl	507	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
435	36.5	22.0	377	10	US-09-849-031A-3	Sequence 813, App	508	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
436	36.5	22.0	377	10	US-09-849-031A-3	Sequence 31, Appl1	509	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
437	36.5	22.0	408	9	US-09-712-363-268	Sequence 13180, A	510	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
438	36.5	22.0	431	10	US-09-815-242-11861	Sequence 13639, A	511	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
439	36.5	22.0	431	10	US-09-815-242-13639	Sequence 10200, A	512	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
440	36.5	22.0	475	10	US-09-815-242-10200	Sequence 13821, A	513	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
441	36.5	22.0	476	10	US-09-815-242-13821	Sequence 6778, Ap	514	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
442	36.5	22.0	498	9	US-09-738-626-6678	Sequence 6778, Ap	515	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
443	36.5	22.0	507	9	US-09-738-626-6732	Sequence 16, Appl	516	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
444	36.5	22.0	525	10	US-09-731-393-16	Sequence 10, Appl	517	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
445	36.5	22.0	527	10	US-09-731-393-10	Sequence 5557, Ap	518	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
446	36.5	22.0	535	10	US-09-815-242-5557	Sequence 12510, A	519	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
447	36.5	22.0	540	10	US-09-815-242-12510	Sequence 8, Appl1	520	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
448	36.5	22.0	545	10	US-09-988-200-8	Sequence 813, App	521	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
449	36.5	22.0	566	10	US-09-943-671-7	Sequence 7, Appl1	522	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
450	36.5	22.0	613	10	US-09-943-671-7	Sequence 31, Appl	523	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
451	36.5	22.0	655	9	US-09-949-842-18	Sequence 329, App	524	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
452	36.5	22.0	655	10	US-09-943-671-31	Sequence 3, Appl1	525	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
453	36.5	22.0	715	10	US-09-841-132-329	Sequence 5, Appl1	526	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
454	36.5	22.0	830	9	US-10-071-503-3	Sequence 5, Appl1	527	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
455	36.5	22.0	831	9	US-10-033-297-5	Sequence 5, Appl1	528	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
456	36.5	22.0	831	9	US-10-081-806-5	Sequence 5, Appl1	529	36	21.7	201	9	US-09-978-697-477	Sequence 477, App
457	36.5	22.0	831	9	US-10-074-328-5	Sequence 5, Appl1	530	36	21.7	201	9	US-09-978-697-477	Sequence 477, App

531	36	21.7	244	9	US-10-121-049-156	Sequence 156, App	604	36	21.7	865	9	US-10-055-364-24	Sequence 24, Appl
532	36	21.7	244	9	US-10-123-904-156	Sequence 156, App	605	36	21.7	996	10	US-09-815-824-5251	Sequence 5251, Ap
533	36	21.7	244	9	US-10-140-470-156	Sequence 156, App	606	36	21.7	1008	12	US-10-017-828-7	Sequence 7, Appl1
534	36	21.7	244	9	US-10-175-746-156	Sequence 156, App	607	36	21.7	1008	12	US-10-017-828-8	Sequence 8, Appl1
535	36	21.7	244	9	US-10-175-746-156	Sequence 156, App	608	36	21.7	1009	10	US-09-815-824-1241	Sequence 1241, A
536	36	21.7	244	9	US-10-175-746-156	Sequence 156, App	609	36	21.7	1193	12	US-10-040-919-2	Sequence 2, Appl1
537	36	21.7	244	10	US-09-731-872-359	Sequence 359, App	610	36	21.7	1239	9	US-10-076-622-577	Sequence 577, App
538	36	21.7	247	12	US-10-076-622-571	Sequence 571, App	611	36	21.7	1239	12	US-10-007-805-577	Sequence 577, App
539	36	21.7	247	12	US-10-007-805-571	Sequence 571, App	612	36	21.7	1341	12	US-10-007-622-565	Sequence 565, App
540	36	21.7	258	10	US-09-815-242-4947	Sequence 4947, Ap	613	36	21.7	1341	12	US-10-007-805-565	Sequence 565, App
541	36	21.7	258	10	US-09-840-704-3	Sequence 3, Appl1	614	36	21.7	1349	9	US-10-076-622-573	Sequence 573, App
542	36	21.7	261	10	US-09-815-242-10487	Sequence 10487, A	615	36	21.7	1349	12	US-10-007-805-573	Sequence 573, App
543	36	21.7	283	12	US-09-738-626-3603	Sequence 3603, Ap	616	36	21.7	1642	10	US-09-925-442-2	Sequence 2, Appl1
544	36	21.7	283	12	US-10-062-254-286	Sequence 286, App	617	36	21.7	1648	10	US-09-925-442-35	Sequence 35, Appl1
545	36	21.7	290	10	US-09-881-752A-270	Sequence 270, App	618	36	21.7	2697	10	US-09-961-557A-5	Sequence 5, Appl1
546	36	21.7	292	10	US-09-886-055-315	Sequence 315, App	619	36	21.7	2972	8	US-08-424-550B-387	Sequence 387, App
547	36	21.7	293	9	US-09-971-536-71	Sequence 71, Appl	620	35.5	21.4	38	10	US-09-925-301-1495	Sequence 1495, Ap
548	36	21.7	298	10	US-09-801-368-66	Sequence 66, Appl	621	35.5	21.4	38	10	US-09-925-300-1746	Sequence 1746, Ap
549	36	21.7	303	9	US-09-854-133-438	Sequence 438, App	622	35.5	21.4	81	10	US-09-764-847-934	Sequence 934, App
550	36	21.7	303	9	US-09-738-973-438	Sequence 5340, Ap	623	35.5	21.4	86	10	US-09-764-860-544	Sequence 544, App
551	36	21.7	321	9	US-09-738-626-5340	Sequence 1, Appl1	624	35.5	21.4	126	10	US-09-815-242-5124	Sequence 5124, Ap
552	36	21.7	336	10	US-09-747-804-1	Sequence 36, Appl	625	35.5	21.4	187	9	US-09-992-598-189	Sequence 189, App
553	36	21.7	339	10	US-09-970-711-36	Sequence 9, Appl1	626	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
554	36	21.7	367	9	US-09-939-633-9	Sequence 9, Appl1	627	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
555	36	21.7	367	10	US-09-939-754-9	Sequence 9, Appl1	628	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
556	36	21.7	367	10	US-09-939-754-9	Sequence 9, Appl1	629	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
557	36	21.7	378	10	US-09-839-832-9	Sequence 9, Appl1	630	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
558	36	21.7	384	9	US-09-815-242-5137	Sequence 5137, Ap	631	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
559	36	21.7	384	9	US-09-924-400-326	Sequence 326, App	632	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
560	36	21.7	384	9	US-09-924-400-332	Sequence 332, App	633	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
561	36	21.7	384	10	US-09-924-400-334	Sequence 334, App	634	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
562	36	21.7	384	10	US-09-810-936-326	Sequence 326, App	635	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
563	36	21.7	384	10	US-09-810-936-332	Sequence 332, App	636	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
564	36	21.7	384	9	US-09-810-936-334	Sequence 334, App	637	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
565	36	21.7	384	9	US-09-810-936-334	Sequence 334, App	638	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
566	36	21.7	395	9	US-09-712-563-164	Sequence 164, App	639	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
567	36	21.7	397	10	US-09-842-987-2	Sequence 2, Appl1	640	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
568	36	21.7	399	9	US-10-076-622-570	Sequence 570, Appl	641	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
569	36	21.7	399	12	US-10-007-805-572	Sequence 572, App	642	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
570	36	21.7	399	12	US-10-007-805-572	Sequence 572, App	643	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
571	36	21.7	426	12	US-10-007-805-572	Sequence 572, App	644	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
572	36	21.7	426	12	US-10-007-805-572	Sequence 572, App	645	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
573	36	21.7	450	9	US-09-977-660-7	Sequence 7, Appl1	646	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
574	36	21.7	450	10	US-09-815-242-5214	Sequence 5214, Ap	647	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
575	36	21.7	450	10	US-09-977-269-7	Sequence 7, Appl1	648	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
576	36	21.7	523	10	US-09-912-020-354	Sequence 354, App	649	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
577	36	21.7	536	10	US-09-741-660-465	Sequence 465, App	650	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
578	36	21.7	550	9	US-10-076-785-32	Sequence 32, Appl	651	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
579	36	21.7	550	9	US-09-738-626-5843	Sequence 5843, Ap	652	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
580	36	21.7	550	10	US-09-943-671-25	Sequence 25, Appl	653	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
581	36	21.7	597	9	US-09-778-819-2	Sequence 2, Appl1	654	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
582	36	21.7	605	9	US-10-028-072-160	Sequence 160, App	655	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
583	36	21.7	605	9	US-10-121-049-160	Sequence 160, App	656	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
584	36	21.7	605	9	US-10-123-904-160	Sequence 160, App	657	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
585	36	21.7	605	9	US-10-140-470-160	Sequence 160, App	658	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
586	36	21.7	605	9	US-09-832-129-37	Sequence 37, Appl	659	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
587	36	21.7	605	9	US-10-173-746-160	Sequence 160, App	660	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
588	36	21.7	605	9	US-10-176-918-160	Sequence 160, App	661	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
589	36	21.7	605	9	US-10-176-921-160	Sequence 160, App	662	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
590	36	21.7	605	10	US-09-860-298-2	Sequence 2, Appl1	663	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
591	36	21.7	683	10	US-09-738-626-6059	Sequence 6059, Ap	664	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
592	36	21.7	683	12	US-09-770-657-2	Sequence 2, Appl1	665	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
593	36	21.7	683	12	US-10-025-676-1	Sequence 1, Appl1	666	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
594	36	21.7	709	10	US-09-815-242-13349	Sequence 13349, A	667	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
595	36	21.7	740	10	US-09-771-161A-265	Sequence 265, App	668	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
596	36	21.7	740	10	US-09-771-161A-265	Sequence 265, App	669	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
597	36	21.7	745	9	US-09-844-988-10	Sequence 10, Appl	670	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
598	36	21.7	745	10	US-09-796-872-2	Sequence 2, Appl1	671	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
599	36	21.7	760	9	US-09-844-908-10	Sequence 10, Appl	672	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
600	36	21.7	760	9	US-09-712-363-292	Sequence 292, Appl	673	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
601	36	21.7	762	10	US-10-076-785-28	Sequence 28, Appl	674	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
602	36	21.7	762	10	US-09-804-969-15	Sequence 15, Appl	675	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
603	36	21.7	792	10	US-09-908-664-2	Sequence 2, Appl1	676	35.5	21.4	187	10	US-09-989-733-189	Sequence 189, App
					US-09-815-242-11880	Sequence 11880, A							

677	35.5	21.4	187	10	US-09-990-442-189	Sequence 189, App	750	35	21.1	166	9	US-10-175-746-8	Sequence 8, Appl1
678	35.5	21.4	187	10	US-09-991-163-189	Sequence 189, App	751	35	21.1	166	9	US-10-176-918-8	Sequence 8, Appl1
679	35.5	21.4	187	10	US-09-993-604-189	Sequence 189, App	752	35	21.1	166	9	US-10-176-921-8	Sequence 8, Appl1
680	35.5	21.4	187	10	US-09-990-456-189	Sequence 189, App	753	35	21.1	171	10	US-09-864-761-43383	Sequence 43383, A
681	35.5	21.4	187	10	US-09-989-721-189	Sequence 189, App	754	35	21.1	176	10	US-09-746-284-12	Sequence 12, Appl1
682	35.5	21.4	187	12	US-10-052-586-426	Sequence 426, App	755	35	21.1	194	10	US-09-772-105-19	Sequence 19, Appl1
683	35.5	21.4	196	10	US-09-925-299-1008	Sequence 1008, App	756	35	21.1	203	9	US-09-738-626-6095	Sequence 6095, App
684	35.5	21.4	252	10	US-09-893-737-52	Sequence 52, Appl1	757	35	21.1	204	9	US-09-949-842-25	Sequence 25, Appl1
685	35.5	21.4	254	10	US-09-815-242-13899	Sequence 13899, A	758	35	21.1	237	10	US-09-925-301-854	Sequence 854, App
686	35.5	21.4	263	9	US-09-924-340-40	Sequence 40, Appl1	759	35	21.1	254	9	US-09-738-626-4681	Sequence 4681, App
687	35.5	21.4	263	9	US-09-992-600A-40	Sequence 40, Appl1	760	35	21.1	264	9	US-09-510-332-48	Sequence 48, Appl1
688	35.5	21.4	266	10	US-09-815-242-10835	Sequence 10835, A	761	35	21.1	266	10	US-09-772-105-12	Sequence 12, Appl1
689	35.5	21.4	290	9	US-09-738-626-4723	Sequence 4723, App	762	35	21.1	285	9	US-09-738-626-5716	Sequence 5716, App
690	35.5	21.4	298	10	US-09-925-297-802	Sequence 802, App	763	35	21.1	286	10	US-09-815-242-13838	Sequence 13838, A
691	35.5	21.4	317	10	US-09-825-561A-84	Sequence 84, Appl1	764	35	21.1	286	10	US-09-815-242-13835	Sequence 13835, A
692	35.5	21.4	321	10	US-09-886-055-259	Sequence 259, App	765	35	21.1	286	10	US-09-943-671-13	Sequence 13, Appl1
693	35.5	21.4	327	10	US-09-870-962-5	Sequence 5, Appl1	766	35	21.1	288	9	US-09-738-626-4503	Sequence 4503, App
694	35.5	21.4	380	9	US-10-104-408-1	Sequence 1, Appl1	767	35	21.1	291	10	US-09-741-669-375	Sequence 375, App
695	35.5	21.4	380	9	US-09-780-926-1	Sequence 1, Appl1	768	35	21.1	297	10	US-09-815-242-11721	Sequence 13721, A
696	35.5	21.4	380	10	US-09-871-617-2	Sequence 2, Appl1	769	35	21.1	303	9	US-09-916-338A-2	Sequence 2, Appl1
697	35.5	21.4	383	10	US-09-764-803A-23	Sequence 23, Appl1	770	35	21.1	303	10	US-09-838-561-13	Sequence 13, Appl1
698	35.5	21.4	404	9	US-09-827-708A-2	Sequence 2, Appl1	771	35	21.1	303	10	US-09-816-760-13	Sequence 13, Appl1
699	35.5	21.4	404	10	US-09-954-697-6	Sequence 6, Appl1	772	35	21.1	328	10	US-09-942-146-4	Sequence 4, Appl1
700	35.5	21.4	404	10	US-09-888-243-30	Sequence 30, Appl1	773	35	21.1	328	10	US-09-795-693-35	Sequence 35, Appl1
701	35.5	21.4	405	9	US-09-738-626-6717	Sequence 6717, App	774	35	21.1	328	10	US-09-938-970-4	Sequence 4, Appl1
702	35.5	21.4	420	10	US-09-946-142-2	Sequence 2, Appl1	775	35	21.1	338	10	US-09-815-242-11750	Sequence 11750, A
703	35.5	21.4	429	9	US-09-854-133-391	Sequence 391, App	776	35	21.1	342	9	US-09-764-864-1146	Sequence 1146, App
704	35.5	21.4	429	10	US-09-835-995-2	Sequence 2, Appl1	777	35	21.1	342	9	US-10-001-857-201	Sequence 201, App
705	35.5	21.4	429	10	US-09-738-626-6181	Sequence 6181, App	778	35	21.1	354	10	US-09-841-132-576	Sequence 576, App
706	35.5	21.4	438	9	US-09-738-626-6181	Sequence 6181, App	779	35	21.1	369	10	US-09-994-427A-9	Sequence 9, Appl1
707	35.5	21.4	447	9	US-09-843-250-23	Sequence 23, Appl1	780	35	21.1	379	9	US-10-047-660-12	Sequence 12, Appl1
708	35.5	21.4	454	10	US-09-815-242-13445	Sequence 13445, A	781	35	21.1	400	12	US-10-055-430-3	Sequence 3, Appl1
709	35.5	21.4	524	10	US-09-815-242-10942	Sequence 10942, A	782	35	21.1	410	10	US-09-963-896-2	Sequence 2, Appl1
710	35.5	21.4	529	12	US-10-011-436-4	Sequence 4, Appl1	783	35	21.1	436	10	US-09-922-501-8	Sequence 8, Appl1
711	35.5	21.4	578	10	US-09-924-358-39	Sequence 39, Appl1	784	35	21.1	449	9	US-09-738-626-6555	Sequence 6555, App
712	35.5	21.4	621	9	US-10-005-211-2	Sequence 2, Appl1	785	35	21.1	477	10	US-09-815-242-11826	Sequence 11826, A
713	35.5	21.4	709	9	US-09-767-870-8	Sequence 8, Appl1	786	35	21.1	485	10	US-09-815-242-11090	Sequence 11090, A
714	35.5	21.4	729	9	US-10-051-902-29	Sequence 29, Appl1	787	35	21.1	485	10	US-09-748-127-4	Sequence 4, Appl1
715	35.5	21.4	729	10	US-10-051-909-29	Sequence 29, Appl1	788	35	21.1	528	9	US-10-174-590-490	Sequence 490, App
716	35.5	21.4	740	10	US-09-897-214-14	Sequence 14, Appl1	789	35	21.1	528	9	US-10-176-758-490	Sequence 490, App
717	35.5	21.4	816	9	US-09-766-720B-2	Sequence 2, Appl1	790	35	21.1	528	9	US-10-175-737-490	Sequence 490, App
718	35.5	21.4	868	9	US-10-041-007-4	Sequence 4, Appl1	791	35	21.1	528	9	US-10-173-706-490	Sequence 490, App
719	35.5	21.4	868	10	US-09-887-586A-56	Sequence 56, Appl1	792	35	21.1	528	9	US-10-175-738-490	Sequence 490, App
720	35.5	21.4	868	10	US-09-903-012-56	Sequence 56, Appl1	793	35	21.1	528	9	US-10-175-852-490	Sequence 490, App
721	35.5	21.4	1316	9	US-09-712-363-177	Sequence 177, App	794	35	21.1	528	9	US-10-176-757-490	Sequence 490, App
722	35	21.1	26	9	US-09-774-639-302	Sequence 302, App	795	35	21.1	528	9	US-10-176-757-490	Sequence 490, App
723	35	21.1	35	10	US-09-894-882-75	Sequence 75, Appl1	796	35	21.1	528	9	US-10-180-552-490	Sequence 490, App
724	35	21.1	56	10	US-09-864-761-45011	Sequence 45011, A	797	35	21.1	528	9	US-10-180-552-490	Sequence 490, App
725	35	21.1	64	10	US-09-864-761-41283	Sequence 41283, A	798	35	21.1	528	9	US-10-180-552-490	Sequence 490, App
726	35	21.1	70	10	US-09-864-761-46514	Sequence 46514, A	799	35	21.1	528	9	US-09-796-753-122	Sequence 122, App
727	35	21.1	75	10	US-09-864-761-44904	Sequence 44904, A	800	35	21.1	528	9	US-10-173-700-490	Sequence 490, App
728	35	21.1	85	10	US-09-864-761-38557	Sequence 38557, A	801	35	21.1	528	9	US-10-174-579-490	Sequence 490, App
729	35	21.1	85	10	US-09-731-872-374	Sequence 374, App	802	35	21.1	528	9	US-10-174-579-490	Sequence 490, App
730	35	21.1	97	10	US-09-764-869-704	Sequence 704, App	803	35	21.1	528	9	US-10-174-582-490	Sequence 490, App
731	35	21.1	109	10	US-09-910-150-32	Sequence 32, Appl1	804	35	21.1	528	9	US-10-174-582-490	Sequence 490, App
732	35	21.1	121	9	US-09-975-719-421	Sequence 421, App	805	35	21.1	528	9	US-10-175-739-490	Sequence 490, App
733	35	21.1	121	10	US-09-764-898-182	Sequence 182, App	806	35	21.1	528	9	US-10-175-740-490	Sequence 490, App
734	35	21.1	122	9	US-10-001-883-81	Sequence 81, Appl1	807	35	21.1	528	9	US-10-175-743-490	Sequence 490, App
735	35	21.1	139	10	US-09-827-949-31	Sequence 31, Appl1	808	35	21.1	528	9	US-10-176-488-490	Sequence 490, App
736	35	21.1	140	10	US-09-764-877-1654	Sequence 1654, App	809	35	21.1	528	9	US-10-176-492-490	Sequence 490, App
737	35	21.1	142	9	US-09-774-639-298	Sequence 298, App	810	35	21.1	528	9	US-10-176-747-490	Sequence 490, App
738	35	21.1	143	9	US-10-013-896-482	Sequence 482, App	811	35	21.1	528	9	US-10-176-750-490	Sequence 490, App
739	35	21.1	143	9	US-09-895-793-482	Sequence 482, App	812	35	21.1	528	9	US-10-176-985-490	Sequence 490, App
740	35	21.1	143	9	US-09-895-814-482	Sequence 482, App	813	35	21.1	528	9	US-10-176-887-490	Sequence 490, App
741	35	21.1	143	9	US-09-759-143-482	Sequence 482, App	814	35	21.1	528	9	US-10-176-991-490	Sequence 490, App
742	35	21.1	143	10	US-09-780-669-482	Sequence 482, App	815	35	21.1	528	9	US-10-176-992-490	Sequence 490, App
743	35	21.1	143	10	US-09-823-827-482	Sequence 482, App	816	35	21.1	528	9	US-10-176-993-490	Sequence 490, App
744	35	21.1	150	9	US-10-147-874-7	Sequence 7, Appl1	817	35	21.1	528	9	US-10-184-658-490	Sequence 490, App
745	35	21.1	152	9	US-09-738-626-5656	Sequence 5656, App	818	35	21.1	528	10	US-09-732-224-2	Sequence 2, Appl1
746	35	21.1	166	9	US-10-028-072-8	Sequence 8, Appl1	819	35	21.1	528	12	US-10-052-586-490	Sequence 490, App
747	35	21.1	166	9	US-10-121-049-8	Sequence 8, Appl1	820	35	21.1	530	9	US-08-738-626-4780	Sequence 4780, App
748	35	21.1	166	9	US-10-123-904-8	Sequence 8, Appl1	821	35	21.1	532	9	US-09-738-626-6268	Sequence 6268, App
749	35	21.1	166	9	US-10-140-470-8	Sequence 8, Appl1	822	35	21.1	543	10	US-09-732-224-5	Sequence 5, Appl1

823	35	21.1	554	10	US-09-746-491-41	Sequence 41, App1	896	34.5	20.8	150	9	US-09-895-793-707	Sequence 707, App
824	35	21.1	631	9	US-10-108-605-169	Sequence 159, App	897	34.5	20.8	150	9	US-09-895-814-707	Sequence 707, App
825	35	21.1	638	10	US-09-815-242-5222	Sequence 5222, Ap	898	34.5	20.8	150	10	US-09-759-143-707	Sequence 707, App
826	35	21.1	642	10	US-09-815-242-12143	Sequence 12143, A	899	34.5	20.8	150	10	US-09-780-659-707	Sequence 707, App
827	35	21.1	649	10	US-09-815-242-5333	Sequence 5333, Ap	900	34.5	20.8	150	10	US-09-822-827-707	Sequence 707, App
828	35	21.1	662	10	US-09-815-242-12343	Sequence 12343, A	901	34.5	20.8	152	10	US-09-994-365-9	Sequence 9, App11
829	35	21.1	663	10	US-09-815-242-10070	Sequence 10070, A	902	34.5	20.8	153	8	US-08-829-402-3	Sequence 3, App11
830	35	21.1	663	10	US-09-815-242-14080	Sequence 14080, A	903	34.5	20.8	158	10	US-09-925-301-1365	Sequence 1365, Ap
831	35	21.1	692	10	US-09-899-471-8	Sequence 8, App11	904	34.5	20.8	179	9	US-10-021-811-4	Sequence 4, App11
832	35	21.1	705	9	US-09-874-503-14	Sequence 14, App1	905	34.5	20.8	200	10	US-09-815-242-10705	Sequence 172, App
833	35	21.1	705	9	US-10-000-157-14	Sequence 162, App	906	34.5	20.8	250	9	US-09-712-363-172	Sequence 5662, Ap
834	35	21.1	705	9	US-10-063-547-162	Sequence 14, App1	907	34.5	20.8	280	10	US-09-738-626-5662	Sequence 4278, Ap
835	35	21.1	705	9	US-09-816-744-14	Sequence 162, App	908	34.5	20.8	301	9	US-09-738-626-4278	Sequence 280, App
836	35	21.1	705	9	US-09-747-359-14	Sequence 58, App	909	34.5	20.8	310	10	US-09-731-872-280	Sequence 3, App11
837	35	21.1	705	9	US-10-174-590-598	Sequence 58, App	910	34.5	20.8	376	10	US-09-783-203-3	Sequence 8, App11
838	35	21.1	705	9	US-10-176-758-598	Sequence 162, App	911	34.5	20.8	376	10	US-09-734-357-4	Sequence 10, App11
839	35	21.1	705	9	US-10-063-502-162	Sequence 58, App	912	34.5	20.8	376	10	US-09-734-357-8	Sequence 12, App11
840	35	21.1	705	9	US-10-173-706-598	Sequence 58, App	913	34.5	20.8	376	10	US-09-734-357-10	Sequence 14013, A
841	35	21.1	705	9	US-10-173-706-598	Sequence 58, App	914	34.5	20.8	422	12	US-10-042-417-4	Sequence 24, App11
842	35	21.1	705	9	US-10-175-738-598	Sequence 58, App	915	34.5	20.8	451	9	US-09-843-250-24	Sequence 6, App11
843	35	21.1	705	9	US-10-175-752-598	Sequence 58, App	916	34.5	20.8	455	10	US-10-108-605-227	Sequence 138, App
844	35	21.1	705	9	US-10-176-482-598	Sequence 58, App	917	34.5	20.8	489	9	US-09-922-598-138	Sequence 138, App
845	35	21.1	705	9	US-10-176-482-598	Sequence 58, App	918	34.5	20.8	489	9	US-09-989-735-138	Sequence 138, App
846	35	21.1	705	9	US-10-176-757-598	Sequence 58, App	919	34.5	20.8	489	9	US-09-990-444-138	Sequence 138, App
847	35	21.1	705	9	US-10-176-913-598	Sequence 58, App	920	34.5	20.8	489	9	US-09-989-730-138	Sequence 138, App
848	35	21.1	705	9	US-10-180-552-598	Sequence 58, App	921	34.5	20.8	489	9	US-09-990-436-138	Sequence 138, App
849	35	21.1	705	9	US-10-180-557-598	Sequence 58, App	922	34.5	20.8	489	9	US-09-991-181-138	Sequence 138, App
850	35	21.1	705	9	US-10-173-700-598	Sequence 58, App	923	34.5	20.8	489	9	US-09-993-687-138	Sequence 138, App
851	35	21.1	705	9	US-10-174-572-598	Sequence 58, App	924	34.5	20.8	489	9	US-09-989-734-138	Sequence 138, App
852	35	21.1	705	9	US-10-174-579-598	Sequence 58, App	925	34.5	20.8	489	9	US-09-990-444-138	Sequence 138, App
853	35	21.1	705	9	US-10-174-582-598	Sequence 58, App	926	34.5	20.8	489	9	US-09-989-730-138	Sequence 138, App
854	35	21.1	705	9	US-10-174-588-598	Sequence 58, App	927	34.5	20.8	489	9	US-09-991-181-138	Sequence 138, App
855	35	21.1	705	9	US-10-175-739-598	Sequence 58, App	928	34.5	20.8	489	9	US-09-993-687-138	Sequence 138, App
856	35	21.1	705	9	US-10-175-740-598	Sequence 58, App	929	34.5	20.8	489	9	US-09-989-734-138	Sequence 138, App
857	35	21.1	705	9	US-10-175-743-598	Sequence 58, App	930	34.5	20.8	489	9	US-09-997-653-138	Sequence 138, App
858	35	21.1	705	9	US-10-176-488-598	Sequence 58, App	931	34.5	20.8	489	9	US-10-174-590-116	Sequence 116, App
859	35	21.1	705	9	US-10-176-492-598	Sequence 58, App	932	34.5	20.8	489	9	US-10-176-758-116	Sequence 116, App
860	35	21.1	705	9	US-10-176-747-598	Sequence 58, App	933	34.5	20.8	489	9	US-10-175-737-116	Sequence 116, App
861	35	21.1	705	9	US-10-176-750-598	Sequence 58, App	934	34.5	20.8	489	9	US-09-993-667-138	Sequence 138, App
862	35	21.1	705	9	US-10-176-985-598	Sequence 58, App	935	34.5	20.8	489	9	US-10-173-706-116	Sequence 116, App
863	35	21.1	705	9	US-10-176-987-598	Sequence 58, App	936	34.5	20.8	489	9	US-10-175-738-116	Sequence 116, App
864	35	21.1	705	9	US-10-176-991-598	Sequence 58, App	937	34.5	20.8	489	9	US-10-175-752-116	Sequence 116, App
865	35	21.1	705	9	US-10-176-992-598	Sequence 58, App	938	34.5	20.8	489	9	US-10-176-482-116	Sequence 116, App
866	35	21.1	705	9	US-10-176-993-598	Sequence 58, App	939	34.5	20.8	489	9	US-10-176-482-116	Sequence 116, App
867	35	21.1	705	9	US-10-184-658-598	Sequence 58, App	940	34.5	20.8	489	9	US-10-180-552-116	Sequence 116, App
868	35	21.1	705	12	US-10-006-867-162	Sequence 162, App	941	34.5	20.8	489	9	US-10-180-552-116	Sequence 116, App
869	35	21.1	705	12	US-10-052-586-598	Sequence 598, App	942	34.5	20.8	489	9	US-10-180-557-116	Sequence 116, App
870	35	21.1	714	10	US-09-529-063-16	Sequence 16, App1	943	34.5	20.8	489	9	US-09-990-562-138	Sequence 138, App
871	35	21.1	718	10	US-09-823-635-2	Sequence 2, App11	944	34.5	20.8	489	9	US-09-997-428-138	Sequence 138, App
872	35	21.1	720	10	US-09-839-408A-13	Sequence 13, App1	945	34.5	20.8	489	9	US-09-997-428-138	Sequence 138, App
873	35	21.1	720	10	US-09-839-408A-30	Sequence 30, App1	946	34.5	20.8	489	9	US-09-997-428-138	Sequence 138, App
874	35	21.1	726	10	US-09-839-408A-21	Sequence 21, App1	947	34.5	20.8	489	9	US-10-173-700-116	Sequence 116, App
875	35	21.1	760	10	US-09-815-242-13994	Sequence 13994, A	948	34.5	20.8	489	9	US-10-174-572-116	Sequence 116, App
876	35	21.1	823	10	US-09-815-242-13348	Sequence 13348, A	949	34.5	20.8	489	9	US-10-174-579-116	Sequence 116, App
877	35	21.1	832	10	US-09-908-322-6	Sequence 6, App11	950	34.5	20.8	489	9	US-10-174-582-116	Sequence 116, App
878	35	21.1	850	10	US-09-985-675-2	Sequence 2, App11	951	34.5	20.8	489	9	US-10-174-588-116	Sequence 116, App
879	35	21.1	856	10	US-09-815-242-11310	Sequence 11310, A	952	34.5	20.8	489	9	US-10-175-739-116	Sequence 116, App
880	35	21.1	856	10	US-09-815-242-11489	Sequence 11489, A	953	34.5	20.8	489	9	US-10-175-740-116	Sequence 116, App
881	35	21.1	874	9	US-09-158-122-6	Sequence 6, App11	954	34.5	20.8	489	9	US-10-175-748-116	Sequence 116, App
882	35	21.1	874	10	US-09-985-675-1	Sequence 1, App11	955	34.5	20.8	489	9	US-10-176-488-116	Sequence 116, App
883	35	21.1	880	10	US-09-223-490-10	Sequence 10, App1	956	34.5	20.8	489	9	US-10-176-492-116	Sequence 116, App
884	35	21.1	948	10	US-09-897-699-2	Sequence 2, App11	957	34.5	20.8	489	9	US-10-176-747-116	Sequence 116, App
885	35	21.1	957	10	US-09-801-574-80	Sequence 80, App1	958	34.5	20.8	489	9	US-10-176-750-116	Sequence 116, App
886	35	21.1	1021	9	US-09-895-913A-340	Sequence 340, App	959	34.5	20.8	489	9	US-10-176-985-116	Sequence 116, App
887	35	21.1	1045	9	US-09-728-626-5941	Sequence 5941, App	960	34.5	20.8	489	9	US-10-176-987-116	Sequence 116, App
888	35	21.1	1050	9	US-10-062-458-4	Sequence 4, App11	961	34.5	20.8	489	9	US-10-176-991-116	Sequence 116, App
889	35	21.1	1382	10	US-09-729-653-2	Sequence 2, App11	962	34.5	20.8	489	9	US-10-176-992-116	Sequence 116, App
890	35	21.1	2710	9	US-10-153-273-12	Sequence 12, App1	963	34.5	20.8	489	9	US-10-176-993-116	Sequence 116, App
891	35	21.1	3503	9	US-10-108-605-237	Sequence 237, App	964	34.5	20.8	489	9	US-10-184-658-116	Sequence 116, App
892	35	21.1	5179	9	US-10-025-380-1068	Sequence 1068, Ap	965	34.5	20.8	489	10	US-09-988-722-138	Sequence 138, App
893	35	21.1	5179	10	US-09-822-217-1068	Sequence 1068, Ap	966	34.5	20.8	489	10	US-09-988-723-138	Sequence 138, App
894	35	21.1	5179	10	US-09-833-263-1068	Sequence 1068, Ap	967	34.5	20.8	489	10	US-09-988-279-138	Sequence 138, App
895	34.5	20.8	150	9	US-10-012-896-707	Sequence 707, App	968	34.5	20.8	489	10	US-09-989-727-138	Sequence 138, App

969 34.5 20.8 489 10 US-09-989-731-138 Sequence 138, App
970 34.5 20.8 489 10 US-09-989-732-138 Sequence 138, App
971 34.5 20.8 489 10 US-09-981-073-138 Sequence 138, App
972 34.5 20.8 489 10 US-09-990-442-138 Sequence 138, App
973 34.5 20.8 489 10 US-09-991-163-138 Sequence 138, App
974 34.5 20.8 489 10 US-09-993-604-138 Sequence 138, App
975 34.5 20.8 489 10 US-09-990-456-138 Sequence 138, App
976 34.5 20.8 489 10 US-09-989-721-138 Sequence 138, App
977 34.5 20.8 489 12 US-10-052-586-116 Sequence 116, App
978 34.5 20.8 513 10 US-09-058-079-15 Sequence 15, App
979 34.5 20.8 523 9 US-10-113-790-1 Sequence 1, Appl
980 34.5 20.8 530 10 US-09-815-242-5570 Sequence 5570, Ap
981 34.5 20.8 530 10 US-09-815-242-12280 Sequence 12280, A
982 34.5 20.8 530 10 US-09-861-451A-18 Sequence 18, Appl
983 34.5 20.8 575 10 US-09-815-242-11275 Sequence 11275, A
984 34.5 20.8 581 9 US-09-738-626-6369 Sequence 6269, Ap
985 34.5 20.8 584 10 US-09-945-825-2 Sequence 2, Appl
986 34.5 20.8 626 10 US-09-771-161A-201 Sequence 201, App
987 34.5 20.8 626 10 US-09-771-161A-202 Sequence 202, App
988 34.5 20.8 626 12 US-10-000-864-6 Sequence 6, Appl
989 34.5 20.8 626 12 US-10-000-864-12 Sequence 12, Appl
990 34.5 20.8 657 10 US-09-815-242-5408 Sequence 5408, Ap
991 34.5 20.8 667 10 US-09-815-242-12490 Sequence 12490, A
992 34.5 20.8 667 10 US-09-815-242-12811 Sequence 12811, A
993 34.5 20.8 734 9 US-10-170-528-2 Sequence 2, Appl
994 34.5 20.8 752 9 US-10-003-295-2 Sequence 2, Appl
995 34.5 20.8 784 10 US-09-905-983-52 Sequence 52, Appl
996 34.5 20.8 822 9 US-10-003-295-4 Sequence 4, Appl
997 34.5 20.8 888 10 US-09-223-490-35 Sequence 35, Appl
998 34.5 20.8 996 9 US-09-738-626-3847 Sequence 3847, Ap
999 34.5 20.8 1160 9 US-10-028-072-234 Sequence 234, App
1000 34.5 20.8 1160 9 US-10-121-049-234 Sequence 234, App

ALIGNMENTS

RESULT 1
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8

Query Match 100.0%; Score 166; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 5.7e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
DB 19 SWRSGSFVQGLFGQHPDVFLMEPAHV 47

RESULT 2
US-10-007-262-1
; Sequence 1, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:

APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-11-12
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 166; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
DB 50 SWRSGSFVQGLFGQHPDVFLMEPAHV 78

RESULT 3
US-09-816-825-2
; Sequence 2, Application US/09816825
; Patent No. US20010051370A1
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CON
; CURRENT APPLICATION NUMBER: US/09/816,825
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/045,284
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-825-2

Query Match 100.0%; Score 166; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
DB 50 SWRSGSFVQGLFGQHPDVFLMEPAHV 78

RESULT 4
US-09-927-602-6
; Sequence 6, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-6

Query Match 96.4%; Score 160; DB 10; Length 169;
Best Local Similarity 96.6%; Pred. No. 4.1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
Db 16 SWRSGSFVGLFGQHPDVFYLMPEAMHV 44

RESULT 5
US-09-927-602-2
Sequence 2, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 395
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-2

Query Match 96.4%; Score 160; DB 10; Length 395;
Best Local Similarity 96.6%; Pred. No. 1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
Db 48 SWRSGSFVGLFGQHPDVFYLMPEAMHV 76

RESULT 6
US-09-927-602-3
Sequence 3, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 395
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
NAME/KEY: VARIANT
LOCATION: (1)...(395)
OTHER INFORMATION: Xaa = any amino acid
US-09-927-602-3

Query Match 96.4%; Score 160; DB 10; Length 395;
Best Local Similarity 96.6%; Pred. No. 1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
Db 48 SWRSGSFVGLFGQHPDVFYLMPEAMHV 76

RESULT 7
US-09-927-602-5
Sequence 5, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-927-602-5

Query Match 96.4%; Score 160; DB 10; Length 418;
Best Local Similarity 96.6%; Pred. No. 1.1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
Db 72 SWRSGSFVGLFGQHPDVFYLMPEAMHV 100

RESULT 8
US-09-927-602-7
Sequence 7, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-7

Query Match 94.6%; Score 157; DB 10; Length 169;
Best Local Similarity 93.1%; Pred. No. 1.1e-15;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
Db 16 SWRSGSFVGLFGQHPDVFYLMPEAMHV 44

RESULT 9

```
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LT 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-4

Query Match          94.6%; Score 157; DB 10; Length 390;
Best Local Similarity 93.1%; Pred. No. 2,7e-15;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYIMEPAWHV 29
DB 49 SWRSGSFFVGLFGQHPDVFYIMEPAWHV 77

RESULT 10
US-09-927-602-9
; Sequence 9, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LT 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-9

Query Match          74.7%; Score 124; DB 10; Length 169;
Best Local Similarity 69.0%; Pred. No. 6,7e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYIMEPAWHV 29
DB 17 TWRSGSFFVGLFGQHPDVFYIMEPAWHV 45

RESULT 11
US-10-212-933-2
; Sequence 2, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
```

```
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1.001AUS
; CURRENT APPLICATION NUMBER: US/10/212,933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-212-933-2

Query Match          74.7%; Score 124; DB 9; Length 483;
Best Local Similarity 69.0%; Pred. No. 2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYIMEPAWHV 29
DB 125 TWRSGSFFVGLFGQHPDVFYIMEPAWHV 153

RESULT 12
US-10-212-933-4
; Sequence 4, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1.001AUS
; CURRENT APPLICATION NUMBER: US/10/212,933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-933-4

Query Match          74.7%; Score 124; DB 9; Length 484;
Best Local Similarity 69.0%; Pred. No. 2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYIMEPAWHV 29
DB 126 TWRSGSFFVGLFGQHPDVFYIMEPAWHV 154

RESULT 13
US-09-833-790-255
; Sequence 255, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
```


APPLICANT: Lodge, Michael J.
APPLICANT: Wang, Tonglong
APPLICANT: Secrist, Heather
APPLICANT: Mohamath, Radooh
APPLICANT: Indrias, Carol Y.
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 255
LENGTH: 531
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-790-255

Query Match 74.7%; Score 124; DB 10; Length 531;
Best Local Similarity 69.0%; Pred. No. 2.2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 SWRSGSFVGOLFQHPDVFYLMPEPAHV 29
DB 173 TWRSGSFGELEFNQNPVFYLPVHV 201

RESULT 14
US-09-927-602-10
Sequence 10, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
TITLE OF INVENTION: Dystrophy
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 179
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-10

Query Match 69.9%; Score 116; DB 10; Length 179;
Best Local Similarity 81.5%; Pred. No. 1e-09;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 RSGSFFVGOLFQHPDVFYLMPEPAHV 29
DB 26 RSGSFFVGOLFQHPDVFYLMPEPAHV 52

RESULT 15
US-09-927-602-11
Sequence 11, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
TITLE OF INVENTION: Dystrophy
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11

NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 174
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-11

Query Match 60.8%; Score 101; DB 10; Length 174;
Best Local Similarity 59.3%; Pred. No. 1.5e-07;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSFFVGOLFQHPDVFYLMPEPAHV 29
DB 22 RTGSSFFGEFFNQNPVFYLPVHV 48

RESULT 16
US-09-738-626-5973
Sequence 5973, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5973
LENGTH: 344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5973

Query Match 30.1%; Score 50; DB 9; Length 344;
Best Local Similarity 47.8%; Pred. No. 7.1;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

OY 8 FVGOLFQHPD-VFYLMPEPAHV 28
DB 187 FYEELKDHDPDVFYLTGAWN 209

RESULT 17
US-09-888-623-2
Sequence 2, Application US/09888623
Patent No. US20020042115A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yongzhong
TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
TITLE OF INVENTION: Synthase Thioesterase-Like Enzyme
FILE REFERENCE: 04974.00510
CURRENT APPLICATION NUMBER: US/09/888,623
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: US 60/214,012
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 60/255,148

```

; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-2

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 265;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 219 AMKDVTSGNKITYQLPGH---FYLLDPA 244

RESULT 18
US-09-888-623-14
; Sequence 14, Application US/09888623
; Patent No. US20020042115A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
; FILE REFERENCE: 04974.00510
; CURRENT APPLICATION NUMBER: US/09/888,623
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,012
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/255,148
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-14

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 265;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 219 AMKDVTSGNKITYQLPGH---FYLLDPA 244

RESULT 19
US-09-888-623-16
; Sequence 16, Application US/09888623
; Patent No. US20020042115A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
; FILE REFERENCE: 04974.00510
; CURRENT APPLICATION NUMBER: US/09/888,623
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,012
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/255,148
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 16
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-16

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 318;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 272 AMKDVTSGNKITYQLPGH---FYLLDPA 297

RESULT 20
US-09-815-242-5057
; Sequence 5057, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trivick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5057
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5057

Query Match
Best Local Similarity 33.3%; Score 48; DB 10; Length 460;
Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 GGSFFVGLFGQHPDYFLMEPA 28
Db 169 GSKFHEMGMLPFAHIDEPWY 192

RESULT 21
US-09-888-615-68
; Sequence 68, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARVOTZAK, GLEN
```

```

1  APPLICANT:  MANNING, GERRAD
2  APPLICANT:  SUPDASNAME, SUCHA
3  TITLE OF INVENTION:  NOVEL PROTEASAS
4  FILE REFERENCE:  036002/1214
5  CURRENT APPLICATION NUMBER:  US/09/888, 615
6  PRIORITY FILING DATE:  2001-06-26
7  PRIOR APPLICATION NUMBER:  60/214, 047
8  PRIOR FILING DATE:  2000-06-26
9  NUMBER OF SEQ ID NOS:  150
10 SOFTWARE:  PatentIn Ver. 2.1
11  SEQ ID NO 68
12  LENGTH:  753
13  TYPE:  PRT
14  ORGANISM:  Homo sapiens
15  OS-09-888-615-68

```

Query Match	28.9%	Score 48	DB 10	Length 753
Best Local Similarity	70.0%	Pred. NO. 32		
Matches 7	Conservative 2	Mismatches 1	Indels 0	Gaps 0

```
QY      19 VFYLMEPAWH 28
          ||| :||
Db      99 VFYLPQPSWH 108
```

RESULT 22
TC-09-915

```

US-09-815,242-12050
? Sequence: 12050, Application US/09815242
? Patent No. US20020061569A1
? GENERAL INFORMATION:
? APPLICANT: Haselbeck, Robert
? APPLICANT: Ohlsen, Karl L.
? APPLICANT: Zyskind, Judith W.
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John D.
? APPLICANT: Carr, Grant J.
? APPLICANT: Yamamoto, Robert T.
? APPLICANT: Xu, H. Howard
? TITLE OF INVENTION: Identification of Essential Genes in
? TITLE OF INVENTION: Prokaryotes
? FILE REFERENCE: E11TRA.011A
? CURRENT APPLICATION NUMBER: US/09/815,242
? CURRENT FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? NUMBER OF SEQ ID NOS: 14110
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 12050
? LENGTH: 701
? TYPE: PRT
? ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12050

```

Query Match	28.3%	Score 47	DB 10	Length 701
Similarity	32.1%	Pred. No. 41		
Best Local	32.1%			
Matches	9	Conservative	9	Indels 4
				Gaps 1

OY 4 SSSSVGQL---FGQHPDVFLMEPAW 27
 | | | : | | : | : | :
Db 140 SGIPFAGPIGARVGFHPEIGILNPITY 167

RESULT 23
US-09-738

```

US-09-738-626-3608
? Sequence 3608, Application US/09738626
? Publication No. US20020197605A1
? GENERAL INFORMATION:
? APPLICANT: NAKAGAWA, SATOSHI
? APPLICANT: MIZOGUCHI, HIROSHI
? APPLICANT: ANDO, SEIKO
? APPLICANT: HAYASHI, MIKIO
? APPLICANT: OCHIAI, KEIKO
? APPLICANT: YOKOI, HARUHIKO
? APPLICANT: TATEISHI, NAOKO
? APPLICANT: SENOH, AKIHIRO
? APPLICANT: IKEDA, MASATO
? APPLICANT: OZAKI, AKIO
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
? FILE REFERENCE: 249-125
? CURRENT APPLICATION NUMBER: US/09/738,626
? CURRENT FILING DATE: 2000-12-18
? PRIOR APPLICATION NUMBER: JP 99/377484
? PRIOR FILING DATE: 1999-12-16
? PRIOR APPLICATION NUMBER: JP 00/159152
? PRIOR FILING DATE: 2000-04-07
? PRIOR APPLICATION NUMBER: JP 00/280988
? PRIOR FILING DATE: 2000-08-03
? NUMBER OF SEQ ID NOS: 7059
? SOFTWARE: PatentIn ver. 3.0
? SEQ ID NO 3608
? LENGTH: 925
? TYPE: prt
? ORGANISM: Corynebacterium glutamicum
US-09-738-626-3608

```

Query Match	28.3%	Score 47	DB 9	Length 925
Best Local Similarity	41.7%	Pred. No. 55		
Matches 10	Conservative 5	Mismatches 9	Indels 0	Gaps 0

QY 3 RSGSSFVGQLEFGQHPDVEYIMEPA 26
| : | | | : | | : |
Db 78 RDPARFPGQGDGSPDLEFWLYKNA 101

RESULT 24
 118-09-738

US-09-738-626-3746
 / Sequence 3746, Application US/09738626
 / Publication No. US20020197605A1
 /
 / GENERAL INFORMATION:
 / APPLICANT: NAKAGAWA, SATOSHI
 / APPLICANT: MIZOGUCHI, HIROSHI
 / APPLICANT: ANDO, SEIKO
 / APPLICANT: HAYASHI, MIKIRO
 / APPLICANT: OCHIAI, KEIKO
 / APPLICANT: YOKOI, HARUHIKO
 / APPLICANT: TATEISHI, NAOKO
 / APPLICANT: SENOH, AKIHIRO
 / APPLICANT: IKEDA, MASATO
 / APPLICANT: OZAKI, AKIO
 /
 / TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
 / FILE REFERENCE: 249-125
 / CURRENT APPLICATION NUMBER: US/09/738,626
 / CURRENT FILING DATE: 2000-12-18
 / PRIOR APPLICATION NUMBER: JP 99/377484
 / PRIOR FILING DATE: 1999-12-16
 / PRIOR APPLICATION NUMBER: JP 00/159162
 / PRIOR FILING DATE: 2000-04-07
 / PRIOR APPLICATION NUMBER: JP 00/280988
 / PRIOR FILING DATE: 2000-08-03
 / NUMBER OF SEQ ID NOS: 7059
 / SOFTWARE: PatentIn ver. 3.0
 / SEQ. ID NO. 3746
 /
 / LENGTH: 302
 /
 / TYPE: PRN

```

; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3745

Query Match          27.7%; Score 46; DB 9; Length 302;
Best Local Similarity 43.8%; Pred. No. 23;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 14 GQHPDVFYLMPEAMV 29
      |||:|:|:|:|:|:|
Db 179 GEHEKRSIVPEPPMH 194

RESULT 25
US-09-272-809-6
; Sequence 6, Application US/09272809
; Patent No. US20020022239A1
; GENERAL INFORMATION:
; APPLICANT: Lagarias, John C.
; TITLE OF INVENTION: Phytofluors as fluorescent labels
; FILE REFERENCE: 2500.118US0
; CURRENT APPLICATION NUMBER: US/09/272,809
; CURRENT FILING DATE: 1998-03-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:cpb6 locus
; OTHER INFORMATION: SLR12112 (ETRI homolog, PAS domain) an 844 aa
US-09-272-809-6

Query Match          27.4%; Score 45.5; DB 10; Length 844;
Best Local Similarity 30.8%; Pred. No. 82;
Matches 8; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

OY 3 RSGSSFYGLFGQHDPVFILMEP-AM 27
      |:|:|:|:|:|:|:|:|:|
Db 321 RESQFYIEKLANYSPQILYILDFIAW 346

RESULT 26
US-09-864-761-41772
; Sequence 41772, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anncmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41772
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015971.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EST_HUMAN HIT: BF218030.1, EVALUE 1.00e-02
US-09-864-761-41772

Query Match          27.1%; Score 45; DB 10; Length 99;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 WRSSSTYGLFGLFH 16
      ||:|:|:|:|:|:|
Db 73 WRSSSTYGLFGLFH 87

RESULT 27
US-09-925-299-834
; Sequence 834, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 834
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-834

Query Match          27.1%; Score 45; DB 10; Length 239;
Best Local Similarity 33.3%; Pred. No. 26;
```

Matches 9; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 SWRSGSFVQLFGQHPDVFYLMPEAM 27
Db 65 SWRSPVEFPGNLIRGEGIVGTLEEVW 91

RESULT 28

US-09-738-626-6189
; Sequence 6189, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIMAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASAO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6189
; LENGTH: 376
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6189

Query Match 27.1%; Score 45; DB 9; Length 376;
Best Local Similarity 47.8%; Pred. No. 41;
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSVFQQLFGQHPDVFYLMPEAM 27
Db 59 GSSOVGMILGYHRAGFSGMAIAW 81

RESULT 29

US-09-735-705-170
; Sequence 170, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 170
; LENGTH: 791
; TYPE: PR1
; ORGANISM: Homo sapien
US-09-735-705-170

Query Match 27.1%; Score 45; DB 10; Length 791;
Best Local Similarity 57.9%; Pred. No. 91;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SGSSF--VQQLFGQHPDVF 20
Db 746 SGGSFSLGVPAHPDVF 764

RESULT 30

US-09-850-716A-170
; Sequence 170, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 170
; LENGTH: 791
; TYPE: PR1
; ORGANISM: Homo sapien
US-09-850-716A-170

Query Match 27.1%; Score 45; DB 10; Length 791;
Best Local Similarity 57.9%; Pred. No. 91;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SGSSF--VQQLFGQHPDVF 20
Db 746 SGGSFSLGVPAHPDVF 764

RESULT 31

US-09-897-778-170
; Sequence 170, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Mainerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 791
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-897-778-170

Query Match 27.1%; Score 45; DB 10; Length 791;


```
; Sequence 357, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-357
```

```
Query Match          27.1%; Score 45; DB 10; Length 920;
Best Local Similarity 57.9%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGGSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      723 SGGSFVILGVAPGHPDVF 741
```

```
RESULT 37
US-09-897-778-357
; Sequence 357, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Mainerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-357
```

```
Query Match          27.1%; Score 45; DB 10; Length 920;
Best Local Similarity 57.9%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGGSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      723 SGGSFVILGVAPGHPDVF 741
```

```
RESULT 38
US-09-974-298-71
; Sequence 71, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
```

```
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 71
; LENGTH: 942
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 2733282CD1
US-09-974-298-71
```

```
Query Match          27.1%; Score 45; DB 9; Length 942;
Best Local Similarity 57.9%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGGSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      745 SGGSFVILGVAPGHPDVF 763
```

```
RESULT 39
US-09-919-172-87
; Sequence 87, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 87
; LENGTH: 942
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 2733282CD1
US-09-919-172-87
```

```
Query Match          27.1%; Score 45; DB 10; Length 942;
Best Local Similarity 57.9%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGGSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      745 SGGSFVILGVAPGHPDVF 763
```

```
RESULT 40
US-09-735-705-161
; Sequence 161, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735.705
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-735-705-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGCHPDVF 20
DB 746 SSSSFVGLGVPAGPHDVF 764

RESULT 41
US-09-850-716A-161
; Sequence 161, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850.716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-850-716A-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGCHPDVF 20
DB 746 SSSSFVGLGVPAGPHDVF 764

RESULT 42
US-09-897-778-161
; Sequence 161, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedyick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897.778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-897-778-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGCHPDVF 20
DB 746 SSSSFVGLGVPAGPHDVF 764

RESULT 43
US-09-815-242-11294
; Sequence 11294, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11294
; LENGTH: 812
; TYPE: PRF
; ORGANISM: Helicobacter pylori
US-09-815-242-11294

Query Match
Best Local Similarity 26.5%; Score 44; DB 10; Length 812;
Matches 9; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

QY 1 SRRSSSFVQQLFGCHPDVYTLMEP 25
DB 220 AMLGLENVVGRT--NPDEFVFKP 242

RESULT 44
US-09-738-626-5852
; Sequence 5852, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
```



```
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280998
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent ver. 3.0
; SEQ ID NO 5852
; LENGTH: 1054
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5852

Query Match          26.2%; Score 43.5; DB 9; Length 1054;
Best Local Similarity 31.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 11; Gaps 1;

QY 1 SWRSGSFVGLFGQHPDVFYIMEPAHW 29
Db 422 SWRSGEPLI-----YMALPSWEV 439

RESULT 45
US-09-925-302-729
; Sequence 729, Application US/09925502
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent ver. 2.0
; SEQ ID NO 729
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-729

Query Match          25.9%; Score 43; DB 10; Length 466;
Best Local Similarity 37.0%; Pred. No. 1e+02;
Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 2;

QY 4 SGSSFVGLFGQHPDVFYIME-PAHW 29
Db 398 AASTFNG---FQRPITYVMSGPAMQL 421

RESULT 46
US-09-815-242-5480
; Sequence 5480, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
```

```
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5480
; LENGTH: 1160
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5480

Query Match          25.9%; Score 43; DB 10; Length 1160;
Best Local Similarity 42.9%; Pred. No. 2.7e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 SGSSFVGLFGQHPDVFYIME 24
Db 332 SGNGFVGSGFIRKYDRELFIE 352

RESULT 47
US-09-815-242-12151
; Sequence 12151, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12151
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12151

```

```

Query Match
Best Local Similarity 42.9%; Score 43; DB 10; Length 1168;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```

```

OY 4 SGGSFVGLFGQHPDYFLMEP 24
DB 332 SGGSFVGLFGQHPDYFLMEP 352

```

```

RESULT 48
US-09-764-870-510
; Sequence 510, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 510
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-510

```

```

Query Match
Best Local Similarity 47.8%; Score 42.5; DB 10; Length 80;
Matches 11; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

```

```

OY 3 RSGSPFVGLFGQHPDYFLMEP 25
DB 31 RSGSPFVGLFGQHPDYFLMEP 50

```

```

RESULT 49
US-10-001-835-184
; Sequence 184, Application US/10001835
; Patent No. US20020160387A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Ptc
; FILE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 184
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-835-184

```

```

Query Match
Best Local Similarity 33.3%; Score 42.5; DB 9; Length 94;
Matches 12; Conservative 2; Mismatches 9; Indels 13; Gaps 2;

```

```

OY 1 SWRSGSFVGLFGQHPDYFLMEP 27
DB 60 SWRSGSFVGLFGQHPDYFLMEP 91

```

```

RESULT 50
US-09-992-598-281
; Sequence 281, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertlisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04

```

PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252

PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 25.6% Score 42.5; DB 9; Length 229;
Best Local Similarity 37.0%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 7; Gaps 1;

QY 2 WSGSSEVQGLF-----GQHPDVY 21
DB 118 WSGDEAPGLFFVDAERVPCRDHDFV 144

Search completed: February 20, 2003, 13:40:48
Job time : 29.2571 secs

is Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:11 ; Search time 10.6286 Seconds
(Without alignments)
75.222 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :
1: A_Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	169	23	ABB81561 Human N-acetylgluc
2	31	100.0	171	23	ABB81560 Human high endothe
3	31	100.0	179	23	ABB81562 Human keratan sulf
4	31	100.0	183	22	ABB81562 Drosophila melanog
5	31	100.0	380	23	AA011274 Human L-selectin s
6	31	100.0	386	20	AAV39918 Human glycosyl sul
7	31	100.0	386	21	AAV79219 Human transferase
8	31	100.0	386	22	AAV93309 Human polyepitide,
9	31	100.0	395	22	AAV72638 Mouse glycosyl sul
10	31	100.0	395	23	AA011275 Murine intestinal-

11	31	100.0	411	19	AAW61100 Keratan sulphate 6
12	31	100.0	418	23	ABB81557 Mouse intestinal N
13	31	100.0	483	20	AAV31656 Mouse N-acetylgluc
14	31	100.0	484	20	AAV31657 Human N-acetylgluc
15	31	100.0	530	22	AAV93367 Human N-acetylgluc
16	31	100.0	531	22	AAV69414 Lung small cell ca
17	31	100.0	649	22	AAV67037 Drosophila melanog
18	30	96.8	363	22	ABB64512 Drosophila melanog
19	30	96.8	562	23	ABB09452 Glucose-3-dehydrog
20	29	93.5	405	23	ABB92113 Herbicidally activ
21	28	90.3	139	23	ABB32335 Human isomerase-11
22	28	90.3	174	23	ABB81563 Human Chondroitin-
23	28	90.3	253	22	ABB55144 Lactococcus lactis
24	28	90.3	306	17	AAV83040 Capsular polysacch
25	28	90.3	306	20	AAV33895 Amino acid sequenc
26	28	90.3	333	23	ABB91753 Herbicidally activ
27	28	90.3	349	21	AAV44207 Arabidopsis thalia
28	28	90.3	349	21	AAV44206 Arabidopsis thalia
29	28	90.3	354	21	AAV44205 Arabidopsis thalia
30	28	90.3	354	21	AAV44205 Arabidopsis thalia
31	28	90.3	354	21	AAV44205 Arabidopsis thalia
32	28	90.3	354	21	AAV44205 Arabidopsis thalia
33	28	90.3	359	21	AAV44205 Arabidopsis thalia
34	28	90.3	359	21	AAV44205 Arabidopsis thalia
35	28	90.3	359	21	AAV44205 Arabidopsis thalia
36	28	90.3	388	20	AAV39919 Hydroxyjasmonic ac
37	28	90.3	458	18	AAV06480 Mouse glycosyl sul
38	28	90.3	458	18	AAV06480 Chick chondroitin
39	28	90.3	479	19	AAV52863 Putative P. abyssal
40	28	90.3	567	15	AAV66031 Glycosaminoglycan
41	28	90.3	612	23	ABB92573 Mammalian MEK Kina
42	28	90.3	626	16	AAV77546 Herbicidally activ
43	28	90.3	626	16	AAV77546 MEK3 protein. Mu
44	28	90.3	626	19	AAV56159 A mitogen-activate
45	28	90.3	626	20	AAV43320 Mitogen ERK kinase
46	28	90.3	626	20	AAV42106 Human MEK3 protei
47	28	90.3	626	20	AAV42109 Murine MEK3 prote
48	28	90.3	626	20	AAV73533 MEK3 protein. Ma
49	28	90.3	626	21	AAV01218 Murine MEK3. Mus
50	28	90.3	626	21	AAV01218 Murine MEK3. Mus
51	28	90.3	793	22	AAV68414 Drosophila melanog
52	28	90.3	943	19	ABB62085 Streptococcus pneu
53	27	87.1	33	12	AAV62703 Human immune/haema
54	27	87.1	60	22	AAV91689 Peptide #2327 enco
55	27	87.1	69	22	ABB29676 Peptide #2363 enco
56	27	87.1	69	22	ABB34856 Peptide #2271 enco
57	27	87.1	69	22	ABB20272 Human brain expres
58	27	87.1	69	22	AAV53662 Human bone marrow
59	27	87.1	69	22	AAV68044 Peptide #2297 enco
60	27	87.1	69	22	AAV15863 Peptide #8523 enco
61	27	87.1	69	22	AAV22089 Peptide #2408 enco
62	27	87.1	69	22	AAV28371 Peptide #2282 enco
63	27	87.1	69	22	AAV03600 Human isomerase-11
64	27	87.1	69	22	AAV37569 Human isomerase-11
65	27	87.1	87	22	ABG15482 Novel human diagno
66	27	87.1	157	22	AAV89846 C glutamylacum prote
67	27	87.1	157	22	AAV79033 Corynebacterium gl
68	27	87.1	167	22	AAV93213 Human polyepitide,
69	27	87.1	169	23	ABB81558 Human corneal N-ac
70	27	87.1	169	23	ABB81559 Human intestinal N
71	27	87.1	175	21	AAV28198 Human utrophin. H
72	27	87.1	239	18	AAV20567 H. pylori cytoplas
73	27	87.1	272	22	AAV85687 Salmonella typhimur
74	27	87.1	283	21	AAV74530 Neisseria meningit
75	27	87.1	295	17	AAV02613 Human prostatic gr
76	27	87.1	295	17	AAV36557 Human GDF-15 pre-p
77	27	87.1	308	16	AAV77097 Human TGF-beta sup
78	27	87.1	308	16	AAV84710 Human embryonal lu
79	27	87.1	308	18	AAV10672 Human TGF-beta-11k
80	27	87.1	308	18	AAV10673 Human TGF-beta-11k
81	27	87.1	308	18	AAV10672 Human TGF-beta-11k
82	27	87.1	308	18	AAV10672 Human TGF-beta-11k
83	27	87.1	308	18	AAV10666 Human TGF-beta-11k

84	27	87.1	308	18	AAW10669	Human TGF-beta-11k	157	26	83.9	148	21	AA651860	Arabidopsis thalia
85	27	87.1	308	18	AAW10670	Human TGF-beta-11k	158	26	83.9	148	23	AAU72540	Arabidopsis cell c
86	27	87.1	308	19	AAW10671	Human TGF-beta-11k	159	26	83.9	155	22	AAU7354	Novel signal trans
87	27	87.1	308	21	AAW48672	GF-2H protein, Ho	160	26	83.9	170	23	AAU72512	Arabidopsis paria
88	27	87.1	308	21	AAW26122	Human TGF-beta sup	161	26	83.9	249	22	ABW69168	Drosophila melanog
89	27	87.1	308	23	ABW08174	Human NAG-1 protei	162	26	83.9	253	23	ABW61509	Iron uptake ABC tr
90	27	87.1	308	23	AAE13538	Human macrophage 1	163	26	83.9	315	22	AAE10324	Human transporter
91	27	87.1	308	23	AAE13539	Human macrophage 1	164	26	83.9	315	22	AAW94269	Human protein sequ
92	27	87.1	316	21	AAW19726	Human TGF-beta 4	165	26	83.9	315	22	AAW94664	Human protein sequ
93	27	87.1	318	18	AAW10663	Human TGF-beta-11k	166	26	83.9	332	22	AAU17660	Novel signal trans
94	27	87.1	325	22	ABW57775	Drosophila melanog	167	26	83.9	346	22	ABW62612	C glutamicum prote
95	27	87.1	390	22	AAW72639	Human glycosyl sul	168	26	83.9	407	22	AAW91787	Drosophila melanog
96	27	87.1	390	23	ABW81556	Human intestinal N	169	26	83.9	417	13	AAW27725	Sequence transcrib
97	27	87.1	395	22	AAW72640	Human glycosyl sul	170	26	83.9	423	21	AAW03064	Malte KIN17 orthol
98	27	87.1	395	23	ABW81554	Human corneal N-ac	171	26	83.9	423	21	AAW03065	Propionibacterium
99	27	87.1	395	23	ABW81555	Consensus N-acetyl	172	26	83.9	424	22	AAU47211	Hericicidally activ
100	27	87.1	395	23	AAE15438	Human drug metabol	173	26	83.9	449	23	ABW93640	Novel human diagno
101	27	87.1	396	22	AAU36948	Staphylococcus aur	174	26	83.9	463	22	ABW28445	Rat islet glucokin
102	27	87.1	396	22	AAU37307	Staphylococcus aur	175	26	83.9	465	18	AAW37438	Rat liver glucokin
103	27	87.1	405	18	AAW20788	H. pylori cytoplas	176	26	83.9	465	18	AAW37439	Rat liver glucokin
104	27	87.1	418	21	AAW41947	Human ORFX ORF1711	177	26	83.9	465	18	AAW37440	Rat liver glucokin
105	27	87.1	426	22	ABW15281	Novel human diagno	178	26	83.9	465	18	AAW37442	Rat liver glucokin
106	27	87.1	431	23	AAU29327	Streptococcus poly	179	26	83.9	542	22	AAW95387	Human protein sequ
107	27	87.1	439	22	AAU34691	E. coli cellular p	180	26	83.9	662	20	AAW34751	C. pneumoniae prot
108	27	87.1	439	22	AAW98252	Escherichia coli p	181	26	83.9	683	22	ABW21469	Novel human diagno
109	27	87.1	476	22	ABW17759	Novel human diagno	182	26	83.9	724	23	AAW22978	Rice MHI protein.
110	27	87.1	488	23	ABW53898	Lactococcus lactis	183	26	83.9	727	22	AAW9285	Human polypeptide
111	27	87.1	499	16	AAW64269	Shabena epsilon-c	184	26	83.9	739	22	ABW67865	Drosophila melanog
112	27	87.1	511	22	AAW41958	Human polypeptide	185	26	83.9	751	22	ABW60916	Drosophila melanog
113	27	87.1	543	23	AAU11262	Novel protein enco	186	26	83.9	753	22	AAW41051	Human polypeptide
114	27	87.1	551	21	AAW52398	Human keratin KERT	187	26	83.9	757	22	ABW58138	Drosophila melanog
115	27	87.1	551	23	AAE20423	Human keratin-2 (K	188	26	83.9	761	22	ABW31470	Novel human diagno
116	27	87.1	562	23	ABW91595	Herbicidally activ	189	26	83.9	838	21	AAW35885	Arabidopsis thalia
117	27	87.1	568	20	AAW35655	Protein involved i	190	26	83.9	843	21	AAW35884	Arabidopsis thalia
118	27	87.1	604	23	AAE18106	Human nucleoside p	191	26	83.9	843	21	AAW35883	Arabidopsis thalia
119	27	87.1	605	22	ABW32651	Escherichia coli p	192	26	83.9	904	23	AAW020516	Protein of APP rel
120	27	87.1	606	21	AAW56855	Semaphorin K1 poly	193	26	83.9	919	18	AAW37440	Hexokinase-glucoki
121	27	87.1	653	22	AAW47250	Human PPI. Homb s	194	26	83.9	919	18	AAW37444	Hexokinase-glucoki
122	27	87.1	666	23	ABW47488	Listeria monocytog	195	26	83.9	919	18	AAW37436	Hexokinase-glucoki
123	27	87.1	671	23	ABW41598	Human ovarian anti	196	26	83.9	919	18	AAW37433	Hexokinase-glucoki
124	27	87.1	706	22	ABW65428	Drosophila melanog	197	26	83.9	1029	14	AAW38862	CC-A. Rattus ratt
125	27	87.1	815	21	AAW19728	Human SEEX Clone A	198	26	83.9	1171	22	AAU32421	Novel human secret
126	27	87.1	815	22	AAU08678	Human FCTR2 polype	199	26	83.9	1174	11	AAW08257	B. thuringiensis to
127	27	87.1	819	22	AAW78501	Human protein SEQ	200	26	83.9	1174	13	AAW82585	Novel toxin expres
128	27	87.1	838	22	AAW79502	Human protein SEQ	201	26	83.9	1174	17	AAW89494	B.t. toxin 81A2.
129	27	87.1	842	21	AAW19727	Human SEEX Clone 4	202	26	83.9	1174	18	AAW09043	Bacillus thuringie
130	27	87.1	842	22	AAW82300	Human follistatin-	203	26	83.9	1348	22	ABW60761	Drosophila melanog
131	27	87.1	851	22	AAW78324	Human protein SEQ	204	26	83.9	1396	20	AAW36871	Protein involved i
132	27	87.1	983	21	AAW70775	Follistatin-relate	205	26	83.9	2464	22	ABW63174	Drosophila melanog
133	27	87.1	983	23	AAE22858	Human Zfsia2 prote	206	26	83.9	2548	20	AAW53761	Human myosin IXa.
134	27	87.1	1119	22	AAW97825	Thermus aquaticus	207	26	80.6	33	21	AAW81628	Cone snail alpha-c
135	27	87.1	1119	23	AAE14716	Chlamydia pneumoni	208	26	80.6	52	21	AAW08204	Arabidopsis thalia
136	27	87.1	1218	20	AAW34680	Human polypeptide	209	26	80.6	55	22	AAW39137	Human colon cancer
137	27	87.1	1266	22	AAW40172	Human polypeptide	210	26	80.6	65	22	AAW39137	Propionibacterium
138	27	87.1	1833	22	ABW71141	Drosophila melanog	211	26	80.6	72	21	AAW08203	Arabidopsis thalia
139	27	87.1	2257	22	ABW61782	Drosophila melanog	212	26	80.6	88	18	AAW28240	Amino acid sequenc
140	27	87.1	3433	18	AAW22017	Utrrophin. Homo sa	213	26	80.6	102	21	AAW07745	Arabidopsis thalia
141	27	87.1	45	22	ABW26868	Peptide #1359 enco	214	26	80.6	102	21	AAW32541	Arabidopsis thalia
142	27	87.1	45	22	ABW3872	Peptide #1378 enco	215	26	80.6	103	21	AAW49433	Arabidopsis thalia
143	27	87.1	45	22	ABW19311	Protein #1310 enco	216	26	80.6	103	22	AAW72397	Rice proline trans
144	27	87.1	45	22	AAW54637	Human brain expres	217	26	80.6	104	21	AAW07414	Arabidopsis thalia
145	27	87.1	45	22	AAW67043	Human bone marrow	218	26	80.6	104	21	AAW32540	Arabidopsis thalia
146	27	87.1	45	22	AAW14903	Peptide #1337 enco	219	26	80.6	104	21	AAW41699	Arabidopsis thalia
147	27	87.1	45	22	AAW27332	Peptide #1369 enco	220	26	80.6	105	21	AAW49432	Arabidopsis thalia
148	27	87.1	45	22	AAW02628	Peptide #1310 enco	221	26	80.6	112	21	AAW52616	Helicobacter pylor
149	27	87.1	130	21	AAW36699	Human peptide enco	222	26	80.6	122	21	AAW07413	Arabidopsis thalia
150	27	87.1	130	21	AAW06258	Arabidopsis thalia	223	26	80.6	122	21	AAW32539	Arabidopsis thalia
151	27	87.1	130	21	AAW51862	Arabidopsis thalia	224	26	80.6	122	21	AAW41698	Arabidopsis thalia
152	27	87.1	135	21	AAW06257	Arabidopsis thalia	225	26	80.6	123	21	AAW49431	Arabidopsis thalia
153	27	87.1	135	21	AAW51861	Arabidopsis thalia	226	26	80.6	128	11	AAW06232	Variable region of
154	27	87.1	144	22	AAU18576	Human lung antigen	227	26	80.6	129	18	AAW20346	H. pylori transmem
155	27	87.1	144	22	AAU17843	Novel human respir	228	26	80.6	133	21	AAW41697	Arabidopsis thalia
156	27	87.1	148	21	AAW06256	Arabidopsis thalia	229	26	80.6	139	22	AAW01483	Human polypeptide

230	25	80.6	142	22	ABR65566	Drosophila melanog	303	25	80.6	515	22	ABG02659	Novel human diagno
231	25	80.6	146	22	AAU04566	Human G-protein co	304	25	80.6	525	22	ABR60745	Drosophila melanog
232	25	80.6	171	22	AAE03594	Wheat tryptophanyl	305	25	80.6	539	22	ABR62067	Drosophila melanog
233	25	80.6	182	20	AAZ56586	Cockroach allergen	306	25	80.6	568	22	ABR65545	Drosophila melanog
234	25	80.6	182	20	AAW88108	A cockroach allergy	307	25	80.6	599	22	ABR87379	Human gene 38 enco
235	25	80.6	185	23	AAZ62799	Eubacterial DNA po	308	25	80.6	599	22	ABG65336	Human albumin fusi
236	25	80.6	196	18	AAZ20596	H. pylori membrane	309	25	80.6	604	21	ABR34841	Gene 42 human secr
237	25	80.6	196	22	AAW72319	Human olfactory re	310	25	80.6	608	23	ABP39010	Staphylococcus epi
238	25	80.6	243	18	AAW09423	Banana polyphenol	311	25	80.6	615	16	AAW77276	ORC3 subunit of Ye
239	25	80.6	248	14	AAAR36280	HME1. Homo sapien	312	25	80.6	626	21	AAW22226	S. cerevisiae orig
240	25	80.6	248	21	AAAG39090	Arabidopsis thaila	313	25	80.6	626	21	AAW42659	Human ORFX ORF423
241	25	80.6	248	21	AAZ58413	14-3-3 sigma amino	314	25	80.6	626	22	AAE21447	Human MAP/ERK kina
242	25	80.6	249	19	AAW60198	Human mammary epi	315	25	80.6	633	22	ABR57813	Drosophila melanog
243	25	80.6	251	22	ABR68580	Drosophila melanog	316	25	80.6	678	21	ABR42123	Human ORFX ORF187
244	25	80.6	274	22	AAU62696	Propionibacterium	317	25	80.6	739	22	ABR63486	Drosophila melanog
245	25	80.6	284	22	AAW96362	Putative P. abyssi	318	25	80.6	767	23	ABR97532	Novel human protei
246	25	80.6	300	22	AAU60924	Propionibacterium	319	25	80.6	787	22	AAW96500	Putative P. abyssi
247	25	80.6	301	21	AAAG39089	Arabidopsis thaila	320	25	80.6	832	22	ABR68340	A human dishevelle
248	25	80.6	305	23	AAAG61514	Iron uptake ABC tr	321	25	80.6	832	23	AAE19053	Drosophila sp. PAR
249	25	80.6	313	22	AAE05861	Pseudomonas stutze	322	25	80.6	883	16	ABR69627	Bacteriophage T7 R
250	25	80.6	316	22	ABR66006	Drosophila melanog	323	25	80.6	883	20	AAV01322	Wild-type T7 RNA p
251	25	80.6	317	19	AAW60997	Streptococcus pneu	324	25	80.6	883	20	AAV01353	Wild-type T7 RNA p
252	25	80.6	321	23	ABR84748	DNA polymerase III	325	25	80.6	883	20	AAV01286	Wild-type RNA poly
253	25	80.6	322	23	ABP38349	Staphylococcus epi	326	25	80.6	883	21	AAE13779	Mutant thermostabl
254	25	80.6	328	22	ABR67466	Drosophila melanog	327	25	80.6	883	22	ABR74395	Modified bacterioph
255	25	80.6	332	22	ABR46741	H. pylori DnaB pro	328	25	80.6	1058	22	ABR68341	A human dishevelle
256	25	80.6	333	23	ABR93585	Herbicideally activ	329	25	80.6	1154	22	ABG21013	Novel human diagno
257	25	80.6	340	23	ABR93581	Herbicideally activ	330	25	80.6	1205	20	AAV33683	P. falciparum 1-de
258	25	80.6	347	22	AAH49723	Hydroxyjaconic ac	331	25	80.6	1205	21	AAW90882	P. falciparum 1-de
259	25	80.6	371	22	AAU35323	Enterococcus faeca	332	25	80.6	1205	21	AAW21863	Isoprenoid related
260	25	80.6	373	22	AAAG89814	C. glutamicum prote	333	25	80.6	1249	17	AAW21572	Ret brain serine e
261	25	80.6	374	22	AAAG83001	S. epidermidis ope	334	25	80.6	1259	21	AAV96744	Ret brain homology
262	25	80.6	377	22	AAU27595	Fusion protein 730	335	25	80.6	1529	21	AAV96744	A. terreus ORF1 es
263	25	80.6	383	21	AAAG09654	Arabidopsis thaila	336	25	80.6	2233	19	AAW48711	HP1V-3 US isolate
264	25	80.6	383	21	AAAG13572	Arabidopsis thaila	337	25	80.6	2233	19	AAW48712	HP1V-3 FRH1 cp45 v
265	25	80.6	383	21	AAAG17445	Arabidopsis thaila	338	25	80.6	2233	19	AAW48713	HP1V-3 Vero cp45 v
266	25	80.6	383	21	AAAG36744	Arabidopsis thaila	339	25	80.6	2233	21	AAW08626	Amino acid sequenc
267	25	80.6	383	21	AAAG45472	Arabidopsis thaila	340	25	77.4	10	21	AAV58729	Paraneoplastic cer
268	25	80.6	383	21	AAAG45500	Arabidopsis thaila	341	25	77.4	18	22	ABR45293	Rabbit albumin-bi
269	25	80.6	390	22	ABR71416	Drosophila melanog	342	25	77.4	24	17	AAK98169	Peptide fragment c
270	25	80.6	391	23	ABR40035	Staphylococcus epi	343	25	77.4	26	19	AAW38669	Streptococcus pneu
271	25	80.6	400	22	ABE03593	Styphen tryptophan	344	25	77.4	30	22	ABR43156	Peptide #10662 enc
272	25	80.6	402	23	ABR53504	Lactococcus lactis	345	25	77.4	30	22	AAW64064	Human brain expres
273	25	80.6	410	22	AAE03592	Corn tryptophanyl-	346	25	77.4	30	22	AAW76884	Human bone marrow
274	25	80.6	418	22	ABR71839	Drosophila melanog	347	25	77.4	30	22	AAW36992	Peptide #11029 enc
275	25	80.6	421	22	ABR67110	Drosophila melanog	348	25	77.4	39	22	ABG46008	Human peptide enco
276	25	80.6	432	22	AAV72396	Corn proline trans	349	25	77.4	39	21	AAK21630	Cone snail alpha-c
277	25	80.6	436	21	AAAG09653	Arabidopsis thaila	350	25	77.4	41	22	AAW90829	Human immune/haema
278	25	80.6	436	21	AAAG36743	Arabidopsis thaila	351	25	77.4	52	22	AAE11098	Peptide #2 related
279	25	80.6	436	21	ABR92008	Herbicideally activ	352	25	77.4	55	22	AAW66036	Amino acid sequenc
280	25	80.6	439	23	ABR92681	Herbicideally activ	353	25	77.4	55	22	AAW55296	Propionibacterium
281	25	80.6	442	21	AAAG13571	Arabidopsis thaila	354	25	77.4	64	23	ABP09760	Human ORFX protei
282	25	80.6	442	21	AAAG17444	Arabidopsis thaila	355	25	77.4	66	22	ABR68290	Drosophila melanog
283	25	80.6	442	21	AAAG45471	Arabidopsis thaila	356	25	77.4	67	22	ABR39299	Peptide #6805 enco
284	25	80.6	442	21	AAAG45499	Arabidopsis thaila	357	25	77.4	67	22	ABR24112	Protein #6111 enco
285	25	80.6	442	23	ABR92053	Herbicideally activ	358	25	77.4	67	22	AAW5967	Human brain expres
286	25	80.6	466	21	AAAG30338	Arabidopsis thaila	359	25	77.4	67	22	AAW72567	Human bone marrow
287	25	80.6	467	21	AAV75499	Neisseria meningit	360	25	77.4	67	22	AAW19626	Peptide #6060 enco
288	25	80.6	467	22	AAU27594	Neisseria meningit	361	25	77.4	67	22	AAW32807	Peptide #6844 enco
289	25	80.6	472	21	AAAG30337	Arabidopsis thaila	362	25	77.4	67	22	ABG43390	Human peptide enco
290	25	80.6	473	23	AAAB82135	Rice proline trans	363	25	77.4	73	22	ABR79554	Corynebacterium gl
291	25	80.6	476	21	AAAB35542	Human cap methyltr	364	25	77.4	75	23	ABP31115	Human ORF88 protei
292	25	80.6	476	21	ABR32245	Human cap methyltr	365	25	77.4	79	22	AAW64277	Human brain expres
293	25	80.6	476	21	ABE13599	Human RNA (guanine	366	25	77.4	79	22	AAW77101	Human bone marrow
294	25	80.6	481	21	AAAG30336	Arabidopsis thaila	367	25	77.4	84	22	AAW74690	Human colon cancer
295	25	80.6	481	23	AAE07510	Herbicideally activ	368	25	77.4	92	22	AAO12390	Human polypeptide
296	25	80.6	481	23	ABR93601	Herbicideally activ	369	25	77.4	93	19	AAW77706	Galactol-1-phosp
297	25	80.6	487	19	AAW57484	Papaya ACC synthas	370	25	77.4	95	22	ABG03315	Novel human diagno
298	25	80.6	488	19	AAW98732	H. pylori GHPO 921	371	25	77.4	98	22	AAE06977	Mouse germ-line hea
299	25	80.6	488	22	AAU35851	Helicobacter pylori	372	25	77.4	100	22	AAO02878	Human polypeptide
300	25	80.6	498	22	AAO82973	Human homologue of	373	25	77.4	104	23	ABP33788	Human kinase-like
301	25	80.6	505	22	ABR71008	Drosophila melanog	374	25	77.4	109	23	ABP40489	Staphylococcus epi
302	25	80.6	512	22	ABR61369	Drosophila melanog	375	25	77.4	112	22	AAU27722	Bacteriophage lamb

376	24	77.4	115	21	AAV90814	33F8 hybridoma VH	449	24	77.4	233	22	AG91894	C glutamicum prote
377	24	77.4	116	17	AAU03642	Anti-adipocyte mon	450	24	77.4	233	23	ABR49753	Listeria monocytog
378	24	77.4	116	22	AAU02515	Drosophila melanog	451	24	77.4	234	14	ABR45009	Sequence encoded b
379	24	77.4	117	22	ABR68972	Novel human diagno	452	24	77.4	234	23	AAU78086	Human CD30-11gand
380	24	77.4	118	22	ABG06010	Novel human diagno	453	24	77.4	236	14	AAU44016	"Lys-63" E.coli he
381	24	77.4	119	18	AAW28320	Amino acid sequenc	454	24	77.4	236	14	AAU44017	"Lys-97" E.coli he
382	24	77.4	119	22	ABG17655	Novel human diagno	455	24	77.4	236	14	AAU44018	"Tyr-97" E.coli he
383	24	77.4	124	23	ABP03481	Human ORFX protein	456	24	77.4	236	14	AAU44019	"Glu-107" E.coli h
384	24	77.4	125	16	AAU80333	Protein polymERIC	457	24	77.4	236	14	AAU44020	"Lys-104" E.coli h
385	24	77.4	128	22	AAU87626	Bovine mammary tis	458	24	77.4	236	14	AAU44021	"Asp-104" E.coli h
386	24	77.4	129	21	AAU56935	Breast and ovarian	459	24	77.4	236	14	AAU44022	"Ser-104" E.coli h
387	24	77.4	129	23	ABU79472	Human TNFR/NGFR 14	460	24	77.4	236	14	AAU44023	"Ser-106" E.coli h
388	24	77.4	130	23	ABP061027	Human secreted pro	461	24	77.4	236	14	AAU44024	"Glu-114" E.coli h
389	24	77.4	130	23	ABP06236	Human ORFX protein	462	24	77.4	236	14	AAU44025	"Lys-114" E.coli h
390	24	77.4	132	21	AAU01026	Human secreted pro	463	24	77.4	236	14	AAU44028	E.coli heat labile
391	24	77.4	133	22	ABG03939	Novel human diagno	464	24	77.4	236	20	AAU92484	Human protein sequ
392	24	77.4	133	22	ABG06007	Novel human diagno	465	24	77.4	237	22	AAU67772	E.coli heat labil
393	24	77.4	138	22	ABU56073	Skin cell protein,	466	24	77.4	237	23	ABU07780	E.coli mutant heat
394	24	77.4	138	22	ABU72273	Murine protein iso	467	24	77.4	238	20	AAU39950	Renilla mulleri gr
395	24	77.4	139	22	AAU80720	Human haematologic	468	24	77.4	238	22	AAU08990	R. reniformis Gree
396	24	77.4	139	22	AAU81194	Human haematologic	469	24	77.4	238	22	AAU83915	Amino acid sequenc
397	24	77.4	139	22	AAU81591	Human haematologic	470	24	77.4	238	22	ABU51820	Renilla reniformis
398	24	77.4	139	22	AAU81830	Human haematologic	471	24	77.4	238	23	AAU13380	Renilla mulleri gr
399	24	77.4	141	21	AAU76072	Human skin cell pr	472	24	77.4	239	22	AAU08991	R. reniformis huma
400	24	77.4	141	22	AAU56011	Skin cell protein,	473	24	77.4	239	22	ABU51821	Humanised Renilla
401	24	77.4	141	23	ABU72211	Human protein isol	474	24	77.4	240	19	AAU65074	E.coli LT-A prote
402	24	77.4	142	23	ABU55522	Lactococcus lactis	475	24	77.4	240	19	AAU65075	E.coli LT-A mutan
403	24	77.4	143	21	AAU08276	Amino acid sequenc	476	24	77.4	240	23	ABU07778	E.coli heat-labile
404	24	77.4	147	23	ABU78638	Rat Osl2 protein	477	24	77.4	242	18	AAU55535	H. pylori ORF 02ep
405	24	77.4	151	20	AAU34489	Porphyromonas ging	478	24	77.4	244	22	AAU04355	Propionibacterium
406	24	77.4	159	20	AAU34364	Porphyromonas ging	479	24	77.4	244	22	AAU50848	Propionibacterium
407	24	77.4	160	19	AAU68807	H. pylori GHP0 132	480	24	77.4	244	22	ABU48557	Listeria monocytog
408	24	77.4	160	23	ABU39482	Staphylococcus epi	481	24	77.4	245	22	ABU16388	Novel human diagno
409	24	77.4	161	22	ABG02526	Novel human diagno	482	24	77.4	245	22	AAU23992	Human EST encoded
410	24	77.4	168	22	ABG03881	Novel human diagno	483	24	77.4	245	22	AAU24000	Human EST encoded
411	24	77.4	172	21	AAU32425	Ubiquinone oxidore	484	24	77.4	246	22	ABG21287	Novel human diagno
412	24	77.4	176	21	AAU45174	Gene 2 human secre	485	24	77.4	247	23	ABU45007	Human Blys binding
413	24	77.4	176	21	AAU45175	Human secreted pro	486	24	77.4	247	23	ABU45195	Human Blys binding
414	24	77.4	177	17	AAU98370	Mycobacterial AhpF	487	24	77.4	249	23	ABU44998	Human Blys binding
415	24	77.4	177	21	ABG23822	Arabisdopsis thalia	488	24	77.4	249	23	ABU45008	Human Blys binding
416	24	77.4	184	22	ABG26613	Novel human diagno	489	24	77.4	249	23	ABU45013	Human Blys binding
417	24	77.4	185	23	ABP60704	Arabisdopsis thalia	490	24	77.4	249	23	ABU45013	Human Blys binding
418	24	77.4	186	21	AAU53624	Human colon cancer	491	24	77.4	249	23	ABU45112	Human Blys binding
419	24	77.4	186	21	AAU21658	Arabisdopsis thalia	492	24	77.4	249	23	ABU45191	Human Blys binding
420	24	77.4	187	22	AAU09736	Human kinase inter	493	24	77.4	249	23	ABU45366	Human Blys binding
421	24	77.4	190	21	AAU09214	Arabisdopsis thalia	494	24	77.4	249	23	ABU45559	Human Blys binding
422	24	77.4	192	16	AAU80345	Protein polymERIC	495	24	77.4	249	23	ABU45560	Human Blys binding
423	24	77.4	193	22	ABG26614	Novel human diagno	496	24	77.4	249	23	ABU45561	Human Blys binding
424	24	77.4	193	22	AAU92454	Human protein sequ	497	24	77.4	249	23	ABU45562	Human Blys binding
425	24	77.4	194	21	AAU09213	Arabisdopsis thalia	498	24	77.4	250	23	ABU45003	Human Blys binding
426	24	77.4	194	23	ABU92772	Herbicideally activ	499	24	77.4	250	23	ABU45148	Human Blys binding
427	24	77.4	198	16	AAU80346	Protein polymERIC	500	24	77.4	250	23	ABU45157	Human Blys binding
428	24	77.4	198	19	AAU49722	Protein polymer ad	501	24	77.4	250	23	ABU45201	Human Blys binding
429	24	77.4	198	19	AAU57672	Collagen-like poly	502	24	77.4	250	23	ABU45524	Human Blys binding
430	24	77.4	200	23	ABP26704	Streptococcus poly	503	24	77.4	250	23	ABU45529	Human Blys binding
431	24	77.4	201	21	AAU16880	Arabisdopsis thalia	504	24	77.4	250	23	ABU45552	Human Blys binding
432	24	77.4	203	23	ABP38851	Staphylococcus epi	505	24	77.4	250	23	ABU45553	Human Blys binding
433	24	77.4	204	21	AAU23821	Arabisdopsis thalia	506	24	77.4	250	23	ABU45554	Human Blys binding
434	24	77.4	209	21	AAU43593	Human cancer assoc	507	24	77.4	250	23	ABU45555	Human Blys binding
435	24	77.4	209	22	AAU99984	Chick limb deforma	508	24	77.4	250	23	ABU45563	Human Blys binding
436	24	77.4	213	21	AAU28194	Human mitosis/kine	509	24	77.4	250	23	ABU45566	Human Blys binding
437	24	77.4	215	14	AAU45007	Sequence encoded b	510	24	77.4	251	19	AAU49735	Protein polymer ad
438	24	77.4	216	18	AAU28267	Staphylococcus aur	511	24	77.4	251	23	ABU44879	Human Blys binding
439	24	77.4	218	23	AAU23588	Chick heat shock c	512	24	77.4	251	23	ABU44880	Human Blys binding
440	24	77.4	218	23	AAU23593	Drosophila pseudoc	513	24	77.4	251	23	ABU44883	Human Blys binding
441	24	77.4	218	23	AAU23594	Cricetulus griseus	514	24	77.4	251	23	ABU44892	Human Blys binding
442	24	77.4	218	23	AAU23596	Drosophila melanog	515	24	77.4	251	23	ABU44895	Human Blys binding
443	24	77.4	218	23	AAU23597	Human heat shock p	516	24	77.4	251	23	ABU44999	Human Blys binding
444	24	77.4	218	23	AAU23598	Chick heat shock p	517	24	77.4	251	23	ABU45000	Human Blys binding
445	24	77.4	218	23	AAU23598	Human Blys binding	518	24	77.4	251	23	ABU45002	Human Blys binding
446	24	77.4	225	21	AAU45121	Rattus species hea	519	24	77.4	251	23	ABU45009	Human Blys binding
447	24	77.4	230	22	AAU76724	Human secreted pro	520	24	77.4	251	23	ABU45110	Human Blys binding
448	24	77.4	231	22	AAU68745	Corynebacterium q1	521	24	77.4	251	23	ABU45111	Human Blys binding

522	24	77.4	251	23	ABP45115	Human Blys binding	595	24	77.4	259	21	AAV96645	Plant-optimized E.
523	24	77.4	251	23	ABP45116	Human Blys binding	596	24	77.4	259	21	AAV96647	Synthetic E. coli
524	24	77.4	251	23	ABP45119	Human Blys binding	597	24	77.4	259	21	AAV96648	Plant-optimized E.
525	24	77.4	251	23	ABP45129	Human Blys binding	598	24	77.4	259	21	AAV96649	Plant-optimized E.
526	24	77.4	251	23	ABP45130	Human Blys binding	599	24	77.4	259	21	AAV96650	Plant-optimized E.
527	24	77.4	251	23	ABP45136	Human Blys binding	600	24	77.4	259	21	AAV96651	Plant-optimized E.
528	24	77.4	251	23	ABP45137	Human Blys binding	601	24	77.4	259	21	AAV96652	Plant-optimized E.
529	24	77.4	251	23	ABP45138	Human Blys binding	602	24	77.4	261	18	AAW25125	Single chain prot
530	24	77.4	251	23	ABP45140	Human Blys binding	603	24	77.4	261	18	AAW21712	RIP fusion protein
531	24	77.4	251	23	ABP45141	Human Blys binding	604	24	77.4	262	19	AAW49734	RIP fusion polymer ad
532	24	77.4	251	23	ABP45146	Human Blys binding	605	24	77.4	265	18	AAW25131	Pro-ribosome inact
533	24	77.4	251	23	ABP45146	Human Blys binding	606	24	77.4	265	18	AAW21718	RIP fusion protein
534	24	77.4	251	23	ABP45151	Human Blys binding	607	24	77.4	269	20	AAW87568	Plasmid pUK21-A Al
535	24	77.4	251	23	ABP45155	Human Blys binding	608	24	77.4	271	13	AAW20993	Protein encoded by
536	24	77.4	251	23	ABP45200	Human Blys binding	609	24	77.4	271	14	AAW35542	Kan gene product f
537	24	77.4	251	23	ABP45205	Human Blys binding	610	24	77.4	271	19	AAW14836	Protein encoded by
538	24	77.4	251	23	ABP45208	Human Blys binding	611	24	77.4	271	18	AAW56697	Kanamycin resistan
539	24	77.4	251	23	ABP45209	Human Blys binding	612	24	77.4	271	20	AAV42547	Plasmid pRZRL1, Ka
540	24	77.4	251	23	ABP45213	Human Blys binding	613	24	77.4	271	20	AAV15385	Plasmid pRZRL1, Ka
541	24	77.4	251	23	ABP45221	Human Blys binding	614	24	77.4	271	20	AAV07249	P. aeruginosa kanam
542	24	77.4	251	23	ABP45362	Human Blys binding	615	24	77.4	271	21	AAV70006	Bacterial Kanamycin
543	24	77.4	251	23	ABP45490	Human Blys binding	616	24	77.4	271	22	AAW85258	Protein product of
544	24	77.4	251	23	ABP45545	Human Blys binding	617	24	77.4	271	22	AAW85256	Protein product of
545	24	77.4	251	23	ABP45580	Human Blys binding	618	24	77.4	271	22	AAW00709	pMG707 vector pr
546	24	77.4	251	23	ABP45701	Human Blys binding	619	24	77.4	271	23	AAE23964	Protein #2 encoded
547	24	77.4	251	23	ABP45821	Human Blys binding	620	24	77.4	271	23	AAE18304	Kanamycin resistan
548	24	77.4	251	23	ABP45826	Human Blys binding	621	24	77.4	273	21	AAW23097	Arabidopsis thalia
549	24	77.4	251	23	ABP45829	Human Blys binding	622	24	77.4	276	18	AAW25124	Maize single chain
550	24	77.4	252	23	ABP44866	Human Blys binding	623	24	77.4	276	18	AAW21711	RIP fusion protein
551	24	77.4	252	23	ABP45005	Human Blys binding	624	24	77.4	276	20	AAV00135	Enterococcus faeca
552	24	77.4	252	23	ABP45010	Human Blys binding	625	24	77.4	276	20	ABP43372	E faecalis EF075 a
553	24	77.4	252	23	ABP45037	Human Blys binding	626	24	77.4	277	23	ABP43372	Lactococcus lactis
554	24	77.4	252	23	ABP45117	Human Blys binding	627	24	77.4	278	17		

668	24	77.4	315	21	AA68460	Sphingomonas sp. C	741	24	77.4	398	23	AA021885	Isoprenoid related
669	24	77.4	316	22	AA67156	M. crenulata K1H2	742	24	77.4	402	13	AA020084	Maize RIP deriv. R
670	24	77.4	317	21	AA610986	M. crenulata hemoc	743	24	77.4	402	18	AAW25129	Pro-rubosome inact
671	24	77.4	317	22	AA671098	M. crenulata K1H2	744	24	77.4	402	18	AAW21716	RIP fusion protein
672	24	77.4	318	22	AA640912	Human polypeptide	745	24	77.4	402	20	AAW25164	Maize K1P protein
673	24	77.4	320	20	AAV39288	Human phosphodiester	746	24	77.4	404	23	AA628391	Streptococcus poly
674	24	77.4	320	21	AA645743	Arabidopsis thalia	747	24	77.4	405	22	AA632963	C glutamicum prote
675	24	77.4	323	21	AA620858	Arabidopsis thalia	748	24	77.4	402	21	AA649547	Arabidopsis thalia
676	24	77.4	323	23	AA661055	Lactobacillus tham	749	24	77.4	412	21	AAW95049	Arabidopsis thalia
677	24	77.4	325	20	AAW97705	Staphylococcus aur	750	24	77.4	412	21	AAW55334	Human protein sequ
678	24	77.4	327	21	AA616879	Arabidopsis thalia	751	24	77.4	415	21	AA651603	Arabidopsis thalia
679	24	77.4	327	22	AA605604	Novel human diago	752	24	77.4	415	21	AA696269	putative p. abysci
680	24	77.4	331	17	AA681570	Transglutaminase.	753	24	77.4	419	22	AA651602	Arabidopsis thalia
681	24	77.4	331	21	AA621657	Arabidopsis thalia	754	24	77.4	420	21	AA634980	Arabidopsis thalia
682	24	77.4	331	23	AA606745	Streptomyces lydic	755	24	77.4	420	22	AA679310	Arabidopsis thalia
683	24	77.4	336	21	AA645742	Arabidopsis thalia	756	24	77.4	420	23	AA608836	Corynebacterium g1
684	24	77.4	336	22	AA602575	Enterococcus typt	757	24	77.4	422	22	AA607536	Corynebacterium g1
685	24	77.4	337	21	AA616878	Arabidopsis thalia	758	24	77.4	422	22	AA621286	Novel human diago
686	24	77.4	338	16	AA680350	Arabidopsis thalia	759	24	77.4	423	21	AA653325	Novel human diago
687	24	77.4	338	16	AA680350	Protein polymeric	760	24	77.4	424	21	AA603063	Arabidopsis thalia
688	24	77.4	340	21	AA620857	Arabidopsis thalia	761	24	77.4	425	22	AAU34976	Maize K1N17 orthol
689	24	77.4	340	21	AA620856	Arabidopsis thalia	762	24	77.4	430	13	AA626173	Enterococcus faeca
690	24	77.4	341	21	AA621656	Arabidopsis thalia	763	24	77.4	430	20	AAW39286	Part of Major yo p
691	24	77.4	346	22	AA606006	Novel human diago	764	24	77.4	430	22	AA665234	Phosphodiesterase
692	24	77.4	347	22	AA636873	Staphylococcus aur	765	24	77.4	431	23	AA697375	Human leukocyte al
693	24	77.4	348	22	AA607777	Novel human diago	766	24	77.4	432	22	AA660787	Novel human protei
694	24	77.4	348	22	AA608483	Arabidopsis thalia	767	24	77.4	433	21	AA632095	Arabidopsis thalia
695	24	77.4	350	21	AA617409	Herbicidally activ	768	24	77.4	433	21	AA671386	Arabidopsis thalia
696	24	77.4	350	23	AA691589	Herbicidally activ	769	24	77.4	435	21	AA640644	Myxococcus acyl Co
697	24	77.4	351	23	AA690981	Herbicidally activ	770	24	77.4	435	21	AA640937	Zea mays protein f
698	24	77.4	351	23	AA693636	Herbicidally activ	771	24	77.4	435	21	AA640937	Zea mays protein f
699	24	77.4	352	22	AA693427	Herbicidally activ	772	24	77.4	436	21	AA640643	Zea mays protein f
700	24	77.4	354	15	AA615384	Commamonas testoste	773	24	77.4	436	21	AA640936	Zea mays protein f
701	24	77.4	356	20	AAW97714	Staphylococcus aur	774	24	77.4	436	22	AA644829	Human protein sequ
702	24	77.4	356	23	AAU75492	S. aureus antigen	775	24	77.4	442	22	AA692558	Herbicidally activ
703	24	77.4	358	22	AA681462	S. epidermidis ope	776	24	77.4	443	21	AA603316	Novel human diago
704	24	77.4	359	19	AAW98324	H. pylori GHP0 141	777	24	77.4	449	23	AA692561	Arabidopsis thalia
705	24	77.4	362	13	AA640558	Staphylococcus epi	778	24	77.4	449	23	AA692561	Herbicidally activ
706	24	77.4	363	18	AAW89789	Staphylococcus aur	779	24	77.4	449	23	AA693684	Herbicidally activ
707	24	77.4	367	17	AA694501	Polyhydroxyacid sy	780	24	77.4	450	23	AA648720	Listeria monocytog
708	24	77.4	368	12	AA611218	Human BPPI/Kuntiz	781	24	77.4	451	22	AA656342	Propionibacterium
709	24	77.4	370	23	AA607785	E coli mutant heat	782	24	77.4	452	23	AA692562	Herbicidally activ
710	24	77.4	373	23	AA607785	E coli heat-labile	783	24	77.4	453	23	AA693638	Herbicidally activ
711	24	77.4	375	23	AA697201	Novel human protei	784	24	77.4	456	21	AA640642	Zea mays protein f
712	24	77.4	378	17	AA609219	SELPOR polymer. S	785	24	77.4	458	21	AA691043	Human carboxypepti
713	24	77.4	378	19	AAW53545	Amnio acid sequenc	786	24	77.4	458	22	AAW23932	Human EST encoded
714	24	77.4	380	21	AAW51889	Plasmid pPT0375 pr	787	24	77.4	458	22	AAW24116	Human EST encoded
715	24	77.4	382	22	AAU00507	E. coli heat-labil	788	24	77.4	460	22	AA633883	Herbicidally activ
716	24	77.4	382	22	AAU00506	E. coli heat-labil	789	24	77.4	463	21	AA630856	Amino acid sequenc
717	24	77.4	382	22	AA693621	Human protein sequ	790	24	77.4	464	14	AA650197	Zea mays protein f
718	24	77.4	385	21	AA625461	Arabidopsis thalia	791	24	77.4	464	23	AA622738	Human wild type g1
719	24	77.4	385	21	AA634981	Arabidopsis thalia	792	24	77.4	464	23	AA622738	Human liver glucoc
720	24	77.4	385	23	AA692563	Herbicidally activ	793	24	77.4	465	14	AA630195	Human wild type g1
721	24	77.4	386	22	AAU28045	Novel human secret	794	24	77.4	465	22	AAU68517	Human novel cytoxi
722	24	77.4	387	20	AAU25163	Maize RIP-5 protei	795	24	77.4	465	14	AA692790	Human protein sequ
723	24	77.4	388	21	AA651604	Arabidopsis thalia.	796	24	77.4	466	20	AA650196	Human wild type g1
724	24	77.4	390	23	AA635303	Human familial com	797	24	77.4	469	21	AA639283	Phosphodiesterase
725	24	77.4	391	20	AAV25619	Murine kin17 prote	798	24	77.4	472	12	AA649135	Arabidopsis thalia
726	24	77.4	391	20	AA660850	Mevalonate pathway	799	24	77.4	472	12	AA649135	Arabidopsis thalia
727	24	77.4	391	23	AAU78132	Human familial com	800	24	77.4	473	12	AA656338	Human cannabinoi
728	24	77.4	391	23	AAU78132	Human VDP-1 prote	801	24	77.4	473	12	AA656338	Non-endorphin hum
729	24	77.4	393	23	AAV23620	A human kin17 prot	802	24	77.4	474	22	AA692673	Rat cannabinoi
730	24	77.4	395	23	AA635302	Mouse HXPLIP1. Ma	803	24	77.4	475	22	AA692673	Mouse ischaemic co
731	24	77.4	395	23	AA678131	Mouse HXPLIP1. Ma	804	24	77.4	478	22	AA663526	Human protein sequ
732	24	77.4	396	15	AAW4926	Regulatory protein	805	24	77.4	479	19	AAU30550	Drosophila melanog
733	24	77.4	396	15	AAW4926	Myclamycin IV hyd	806	24	77.4	479	19	AAW49739	Novel human secret
734	24	77.4	397	21	AA633326	Arabidopsis thalia	807	24	77.4	482	23	AA654622	Protein polymer ad
735	24	77.4	398	20	AA62838	Escherichia coli p	808	24	77.4	482	23	AA654622	Lactococcus lactis
736	24	77.4	398	21	AA658134	Lung cancer associ	809	24	77.4	487	21	AA610901	D. melanogaster pe
737	24	77.4	398	21	AA633096	Arabidopsis thalia	810	24	77.4	487	21	AA633456	S. xylois D1A pr
738	24	77.4	398	22	AAU37784	Streptococcus pneu	811	24	77.4	487	21	AAW56734	Human PRO1917 prot
739	24	77.4	398	22	AAU01066	CPE 69 protein seq	812	24	77.4	487	22	AAU12435	Human PRO1917 poly
740	24	77.4	398	23	AAO21878	Isoprenoid related	813	24	77.4	487	22	AA650929	Human PRO1917 prot
												AA64383	Amino acid sequenc

814	24	77.4	487	23	ABB95547	Human angiogenesis	887	24	77.4	635	21	AAB24787	Plant SPD encoded
815	24	77.4	487	23	ABB84941	Human PRO1917 prot	888	24	77.4	636	21	AAG40601	Arabidopsis thalia
816	24	77.4	487	23	AAU083693	Human PRO protein,	889	24	77.4	637	21	ABB55861	Lactococcus lactis
817	24	77.4	489	13	AAAR20085	Matze RIP deriv. R	890	24	77.4	638	23	ABB447300	Listeria monocytog
818	24	77.4	489	18	AAW25130	Pro-ribosome inact	891	24	77.4	639	21	AAB24786	Plant SPD encoded
819	24	77.4	490	18	AAW21717	RIP fusion protein	892	24	77.4	640	16	AAAR32249	Chloroperoxidase.
820	24	77.4	490	24	AAW13366	Amino acid sequenc	893	24	77.4	647	22	ABG00626	Novel human diagno
821	24	77.4	490	22	AAU12326	Human PRO222 polyp	894	24	77.4	652	22	ABB51288	Drosophila melanog
822	24	77.4	490	22	AAAB80234	Human PRO222 prote	895	24	77.4	657	21	AAAG42584	Arabidopsis thalia
823	24	77.4	492	23	ABG65035	Human albumin fusi	896	24	77.4	674	21	AAAB42482	Human ORFX ORF2246
824	24	77.4	492	23	AAU91358	Human secreted pro	897	24	77.4	674	21	AAAG42583	Arabidopsis thalia
825	24	77.4	492	23	AAU74621	Oestrogen-regulate	898	24	77.4	674	21	AAAB03758	Human H37 amino ac
826	24	77.4	493	18	AAW25132	Pro-ribosome inact	899	24	77.4	679	24	ABG10004	Novel human diagno
827	24	77.4	493	18	AAW21719	RIP fusion protein	900	24	77.4	680	23	AAAG6069	N. meningitidis re
828	24	77.4	494	22	ABG62540	Drosophila melanog	901	24	77.4	681	21	AAAG31900	Arabidopsis thalia
829	24	77.4	501	22	ABG58136	Drosophila melanog	902	24	77.4	682	16	AAAR80330	Protein polymeric
830	24	77.4	502	18	AAW25133	Pro-ribosome inact	903	24	77.4	682	19	AAW45717	Protein polymer ad
831	24	77.4	502	18	AAW21720	RIP fusion protein	904	24	77.4	682	19	AAW57670	Collagen-like poly
832	24	77.4	508	22	ABG59820	Drosophila melanog	905	24	77.4	683	21	AAAB24850	Plant SPD encoded
833	24	77.4	509	13	AAAR30429	Major yo paraneopl	906	24	77.4	683	21	AAAG31899	Arabidopsis thalia
834	24	77.4	514	20	AAW33675	Human DKC1 protein	907	24	77.4	683	22	ABR11820	Human secreted pro
835	24	77.4	514	22	AAAB90586	Human secreted pro	908	24	77.4	685	22	ABG58231	Drosophila melanog
836	24	77.4	517	22	ABG65472	Human albumin fusi	909	24	77.4	688	22	ABBB67404	Drosophila melanog
837	24	77.4	517	22	AAE05394	Staphylococcus aur	910	24	77.4	688	22	ABBB93318	Herbicideally activ
838	24	77.4	520	22	AAU34239	Staphylococcus aur	911	24	77.4	689	22	ABBB59166	Drosophila melanog
839	24	77.4	520	22	AAAG90708	C glutamicum prote	912	24	77.4	690	22	ABBB6052	Drosophila melanog
840	24	77.4	520	22	AAAB80105	Cornebacterium gl	913	24	77.4	693	22	AAAG98345	Escherichia coli p
841	24	77.4	521	22	AAE05392	Staphylococcus aur	914	24	77.4	698	21	AAAB24849	Plant SPD encoded
842	24	77.4	521	23	AAW50327	Human nucleotide b	915	24	77.4	703	21	AAAB24848	Plant SPD encoded
843	24	77.4	531	20	AAU05376	Human HCW Inducib	916	24	77.4	703	22	ABBB61081	Drosophila melanog
844	24	77.4	532	17	AAW91035	Recombinant V8 pro	917	24	77.4	717	22	ABBB57943	Drosophila Hsp83 p
845	24	77.4	532	18	AAW22219	Protein encoded by	918	24	77.4	717	22	AAAB36505	Human Hsp90 family
846	24	77.4	533	20	AAW39284	Phosphodiesterase	919	24	77.4	724	22	AAE12989	Human heat shock p
847	24	77.4	533	22	AAAB92829	Human protein sequ	920	24	77.4	724	22	AAAB82537	Human heat shock p
848	24	77.4	536	21	AAAG49134	Arabidopsis thalia	921	24	77.4	724	22	AAAB36507	Human Hsp90 beta p
849	24	77.4	537	18	AAW22220	Protein encoded by	922	24	77.4	724	23	ABBB06994	Human Hsp90 beta p
850	24	77.4	537	22	AAAB60496	Human cell cycle a	923	24	77.4	728	12	AAW09381	Chicken heat shock
851	24	77.4	537	23	AAU74622	Oestrogen-regulate	924	24	77.4	732	17	AAAR10426	Human testicular a
852	24	77.4	540	22	AAU30977	Novel human secret	925	24	77.4	732	21	AAAG42582	Arabidopsis thalia
853	24	77.4	544	20	AAV14049	G. oxydans D-sorbi	926	24	77.4	732	22	AAE12988	Human Hsp90 family
854	24	77.4	555	16	AAAR7864	S. clavuligerus OR	927	24	77.4	732	22	AAAB82536	Human heat shock p
855	24	77.4	556	22	AAE07912	S. clavuligerus cl	928	24	77.4	732	23	AAU71814	Human antigen Huhs
856	24	77.4	559	22	ABG26803	Novel human diagno	929	24	77.4	737	22	AAAM39446	Human polypeptide
857	24	77.4	561	22	ABBS8264	Drosophila melanog	930	24	77.4	738	22	ABG21289	Novel human diagno
858	24	77.4	561	22	AAU38523	Drosophila G-prote	931	24	77.4	761	22	AAAM39447	Human polypeptide
859	24	77.4	562	23	AAU33507	Enterococcus faeca	932	24	77.4	761	23	ABBB84252	Human macro protei
860	24	77.4	569	23	ABBS0001	Listeria monocytog	933	24	77.4	762	16	AAAR80327	Protein polymeric
861	24	77.4	572	22	AAU35140	Enterococcus faeca	934	24	77.4	762	16	AAAR80324	Protein polymeric
862	24	77.4	580	17	AAW02164	Lactococcus lactis	935	24	77.4	762	19	AAW48713	Protein polymer ad
863	24	77.4	580	20	AAW39285	Phosphodiesterase	936	24	77.4	762	19	AAW48715	Protein polymer ad
864	24	77.4	587	21	AAAG49546	Arabidopsis thalia	937	24	77.4	762	19	AAW57666	Collagen-like poly
865	24	77.4	593	20	AAW28561	Cyclic-GMP specifi	938	24	77.4	762	19	AAW57668	Collagen-like poly
866	24	77.4	593	23	ABG61846	Prostate cancer-as	939	24	77.4	762	19	AAW53535	Amino acid sequenc
867	24	77.4	595	22	AAU49951	Propionibacterium	940	24	77.4	770	22	ABG05609	Novel human diagno
868	24	77.4	600	21	AAAG49545	Arabidopsis thalia	941	24	77.4	785	13	ABBB97543	Novel human protei
869	24	77.4	600	22	AAU64743	Propionibacterium	942	24	77.4	789	12	AAAR15143	Hemomonchus contort
870	24	77.4	603	22	AAAG82395	S. epidermidis ope	943	24	77.4	797	18	AAW36508	Human RENT1 protei
871	24	77.4	603	23	ABP40381	Staphylococcus epi	944	24	77.4	798	22	AAW41233	Human polypeptide
872	24	77.4	606	23	AAO17858	Pyth domain conta	945	24	77.4	798	22	AAW41234	Human polypeptide
873	24	77.4	607	21	AAAG53324	Arabidopsis thalia	946	24	77.4	801	21	AAU09864	Novel human secret
874	24	77.4	608	22	AAAB99139	Pyrococcus abyssi	947	24	77.4	807	21	AAAG31898	Arabidopsis thalia
875	24	77.4	608	22	AAU16168	Human novel secret	948	24	77.4	814	22	ABBB5317	Drosophila melanog
876	24	77.4	611	22	AAE12809	Soybean arginyl-tr	949	24	77.4	828	22	ABBB5774	Drosophila melanog
877	24	77.4	613	18	AAW22051	DNA polymerase 220	950	24	77.4	828	22	AAU03139	Streptococcus pyog
878	24	77.4	613	21	AAW97097	Ptu DNA polymerase	951	24	77.4	878	22	ABP26368	Streptococcus poly
879	24	77.4	613	21	AAW52020	P. fultous PFUORF	952	24	77.4	884	16	AAAR80341	Protein polymer ad
880	24	77.4	613	21	AAW51649	Pyrococcus furiosu	953	24	77.4	884	17	AAW09213	SEIP8K polymer. s
881	24	77.4	615	17	AAW00456	1-Fructan fructan	954	24	77.4	884	19	AAW49728	SEIP8K polymer. s
882	24	77.4	616	21	AAAG40603	Arabidopsis thalia	955	24	77.4	884	21	AAW53541	Expected amino aci
883	24	77.4	617	21	AAH24788	Plant SPD encoded	956	24	77.4	884	21	AAW51882	Plasmdid pRO345 pr
884	24	77.4	618	21	AAAG40602	Arabidopsis thalia	957	24	77.4	896	22	AAU61465	Propionibacterium
885	24	77.4	629	22	ABBB73229	Human phosphatase	958	24	77.4	913	22	AAU33722	Pseudomonas aerugi
886	24	77.4	631	22	ABBB58778	Drosophila melanog	959	24	77.4	923	22	ABG13717	Novel human diagno

CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human high endothelial cell
 CC N-acetylglucosamine-6-sulfotransferase (GlcNAc6ST), which is given in
 CC the exemplification of the present invention.

CC Sequence 171 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 23; Length 171;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 129 VRYEDL 134

RESULT 3
 ABB81562
 ID ABB81562 standard; Protein; 179 AA.

XX ABB81562;

DT 05-SEP-2002 (first entry)

DE Human keratan sulfate Gal-6-sulfotransferase portion SEQ ID NO:10.

KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme: GlcNAc6ST;
 KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.

XX Homo sapiens.

PN US2002061562-A1.

PD 23-MAY-2002.

PF 09-AUG-2001; 2001US-0927602.

PR 11-AUG-2000; 2000US-325773P.

PA (FUKUDA/) FUKUDA M N.

PI (AKAMA/) AKAMA T O.

PI Fukuda MN, Akama TO;

DR WPI; 2002-507643/54.

XX New nucleic acid encoding corneal

PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -

PS Example 2; Fig 3C; 69pp; English.

XX The present invention describes human corneal

CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human keratan sulfate
 CC Gal-6-sulfotransferase, which is given in the exemplification
 CC of the present invention.

XX Sequence 179 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 23; Length 179;

XX Sequence 179 AA;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 137 VRYEDL 142

RESULT 4
 ABB68582
 ID ABB68582 standard; Protein; 183 AA.

XX ABB68582;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 32538.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL12685.

XX

PS Disclosure; SEQ ID NO 32538; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 183 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 22; Length 183;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 75 VRYEDL 80

XX Sequence 183 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 22; Length 183;

XX Sequence 183 AA;

XX Sequence 183 AA;

XX Sequence 183 AA;

DT 12-MAR-2002 (first entry)
 XX
 DE Human L-selectin sulfotransferase-2 (LSSR-2) protein.
 XX
 KW Human; beta1,3gnt; beta1,3-N-acetylglucosaminyltransferase; MECA-79;
 KW L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;
 KW ulcerative colitis; inflammatory skin disorder; psoriasis; lichen planus;
 KW allergic contact dermatitis; lymphoma; chronic pneumonia; LSSR-2;
 KW delayed-type hypersensitivity reaction; hyperplastic thymus; antileuker;
 KW antinflammatory; antipsoriatic; antidiabetic; dermatological;
 KW antiallergic.
 XX
 OS Homo sapiens.
 XX
 PN WO200185177-A1.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US15452.
 XX
 PR 11-MAY-2000; 2000US-0569320.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Fukuda M, Yeh J, Hiraoka N;
 XX
 DR WPI; 2002-075226/10.
 XX
 DR N-PSDB; AAS16947.
 XX
 PT New enzyme, useful for modifying acceptor molecule, comprises an
 PT isolated L-selectin sulfotransferase-2 that directs expression of
 PT L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or
 PT intestinal G10NMc 6-sulfotransferase
 XX
 PS Claim 21; Fig 4; 98pp; English.
 XX
 CC The present invention provides a method of modifying an acceptor molecule
 CC by contacting the acceptor with an isolated
 CC beta1,3-N-acetylglucosaminyltransferase (beta1,3gnt) or an active
 CC fragment, where beta1,3gnt directs expression of a MECA-79 antigen. The
 CC invention also provides a method of treating or preventing an
 CC L-selectin-mediated condition by reducing the expression or activity of a
 CC beta1,3gnt that directs expression of a MECA-79 antigen. This can be done
 CC by administering to the subject an oligosaccharide L-selectin antagonist
 CC that inhibits binding of L-selectin to a MECA-79 antigen, for example by
 CC administering antibody material that specifically binds beta1,3gnt,
 CC and/or a beta1,3gnt antisense nucleic acid molecule. L-selectin
 CC sulfotransferase-2 (LSSR-2) also directs MECA-79 antigen expression.
 CC Alternatively, the expression or activity of LSSR-2 or its active
 CC fragment can be reduced in combination with reducing the expression or
 CC activity of beta1,3gnt. The method is useful for treating L-selectin
 CC mediated conditions such as Crohn's disease and ulcerative colitis,
 CC inflammatory disorders of the skin such as allergic contact dermatitis,
 CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type
 CC hypersensitivity reactions, diabetes and hyperplastic thymus. This
 CC sequence represents human LSSR-2.
 XX
 SQ Sequence 380 AA;
 XX
 QY Query Match 100.0%; Score 31; DB 23; Length 380;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VREDEL 6
 XX
 DB 268 VREDEL 273
 XX
 RESULT 6
 XX AAY39918
 XX ID AAY39918 standard; Protein; 386 AA.
 XX
 AC AAY39918;

XX 08-DEC-1999 (first entry)
 DT
 XX
 DE Human glycosyl sulfotransferase-3 protein sequence.
 XX
 KW Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
 KW selectin binding interaction; inflammation; lymphocyte homing; human;
 KW secondary lymph organ.
 XX
 OS Homo sapiens.
 XX
 PN WO9949018-A1.
 XX
 PD 30-SEP-1999.
 XX
 PE 26-FEB-1999; 99WO-US04316.
 XX
 PR 20-MAR-1998; 98US-0045284.
 XX 12-NOV-1998; 98US-0190911.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX (SYNT) SYNTEX USA INC.
 XX
 PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
 XX
 DR WPI; 1999-580442/49.
 XX
 DR N-PSDB; AAZ20792.
 XX
 PT Human and murine glycosyl sulfotransferase 3 and related
 PT polynucleotides
 XX
 PS Claim 2; Fig 1; 59pp; English.
 XX
 CC This sequence is the human glycosyl sulfotransferase-3 (GST-3) of
 CC the invention. The nucleic acid sequences, probes and primers derived
 CC from these, proteins and antibodies are useful in detecting homologues.
 CC The sequences, antibodies and methods are useful in the diagnosis and
 CC treatment of diseases associated with selectin binding interactions;
 CC including conditions associated with or resulting from the homing of
 CC leukocytes to sites of inflammation and the normal homing of lymphocytes
 CC to secondary lymph organs.
 XX
 SQ Sequence 386 AA;
 XX
 QY Query Match 100.0%; Score 31; DB 20; Length 386;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VREDEL 6
 XX
 DB 274 VREDEL 279
 XX
 RESULT 7
 XX AAY79219
 XX ID AAY79219 standard; Protein; 386 AA.
 XX
 AC AAY79219;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Human transferase TRNSFS-11.
 XX
 KW Transferase; TRNSFS-11; human; antitumor; cell proliferation;
 KW gastrointestinal disorder; developmental disorder;
 KW genetic disorder; neurological disorder; reproductive disorder;
 KW smooth muscle disorder; immunological disorder; inflammation;
 KW diagnosis; therapy; N-acetylglucosamine 6-O-sulfotransferase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 121

FT Modified-site /note= "potential O-phosphorylation"
 FT 107
 FT Modified-site /note= "potential O-phosphorylation"
 FT 217
 FT Modified-site /note= "potential O-phosphorylation"
 FT 252
 FT Modified-site /note= "potential O-phosphorylation"
 FT 364
 FT Modified-site /note= "potential O-phosphorylation"
 FT 380
 FT Modified-site /note= "potential O-phosphorylation"
 FT 35
 FT Modified-site /note= "potential O-phosphorylation"
 FT 50
 FT Modified-site /note= "potential O-phosphorylation"
 FT 81
 FT Modified-site /note= "potential O-phosphorylation"
 FT 287
 FT Modified-site /note= "potential O-phosphorylation"
 FT 243
 FT Modified-site /note= "potential O-phosphorylation"
 FT 30
 FT Modified-site /note= "potential N-glycosylation"
 FT 308
 FT Modified-site /note= "potential N-glycosylation"
 FT 329
 FT Modified-site /note= "potential N-glycosylation"
 FT 7...23
 FT Domain /note= "transmembrane domain"
 FT /note= "transmembrane domain"
 PN W0200014251-A2.
 XX 16-MAR-2000.
 PD 09-SEP-1999; 99WO-US20989.
 XX 10-SEP-1998; 98US-0150657.
 PR 04-NOV-1998; 98US-0186779.
 PR 11-MAY-1999; 98US-0133642.
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
 PI Hillman JL, Azimzai Y;
 DR N-PSDB; AA294211.
 DR WPI; 2000-256996/22.
 XX
 PT Human transferase proteins useful for preventing, diagnosing and
 PT treating cancers and developmental, gastrointestinal, genetic,
 PT immunological, neurological, reproductive and smooth muscle disorders -
 XX
 PS Claim 1; Page 90-91; 113pp; English.
 XX
 CC The present sequence is that of human transferase TRNSF-11, 1 of
 CC 15 claimed novel human transferase proteins of the invention (see
 CC AA979209-23). The sequence was deduced from a cDNA clone (see
 CC AA294211) isolated from a galbladder library. It shows homology to
 CC mouse N-acetylglucosamine 6-O-sulfotransferase. TRNSF-11 is
 CC expressed in dermatologic and gastrointestinal tissues,
 CC especially those associated with inflammation and cell
 CC proliferation. The new human transferases and polynucleotides can
 CC be used in the diagnosis, prevention and treatment of cancer,
 CC developmental disorders, gastrointestinal disorders, genetic
 CC disorders, immunological disorders, neurological disorders,
 CC reproductive disorders, and smooth muscle disorders. The
 CC polypeptides can also be used to raise antibodies, and to screen
 CC for agonists and antagonists of transferase activity.
 CC
 SQ Sequence 386 AA;

Query Match 100.0%; Score 31; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 Db 274 VRYEDL 279
 RESULT 8
 AAM93309
 ID AAM93309 standard; Protein: 386 AA.
 AC AAM93309;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide; SEQ ID NO: 2817.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 98JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK944229.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 2817; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 386 AA;

Query Match 100.0%; Score 31; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 Db 274 VRYEDL 279
 RESULT 9
 AAY72638
 ID AAY72638 standard; Protein: 395 AA.
 AC AAY72638;

[illegible]

AC AAW61100;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Keratan sulphate 6-sulphotransferase.
 XX
 KM KSGal6ST; keratan sulphate 6-sulphotransferase;
 KM chick chondroitin 6-sulphotransferase; CGST; phage lambda; hybridization;
 KM expression vector; COS-7 cells; Bluescript plasmid; galactose;
 KM keratan sulphate; chondroitin.
 XX
 OS Homo sapiens.
 XX
 PN EP845533-A2.
 XX
 PD 03-JUN-1998.
 XX
 PF 27-NOV-1997; 97EP-0309564.
 XX
 PR 29-NOV-1996; 96JP-0320535.
 XX
 PA (SECK) SEIKAGAKU KOGYO CO LTD.
 XX
 PI Fukuta M, Habuchi O;
 XX
 DR WPI; 1998-288750/26.
 DR N-PSDB; AAV36418.
 XX
 PT Keratan sulphate 6-sulpho-transferase polypeptide - transfers
 PT sulphate from sulphate donor to galactose 6-hydroxy group etc.
 XX
 PS Claim 4; Page 15-16; 21pp; English.
 XX
 CC This sequence produces the protein KSGal6ST (Keratan sulphate
 CC 6-sulphotransferase), which has a molecular weight of about 46700. The
 CC cDNA sequence of this protein was obtained by radiolabelling the cDNA of
 CC chick chondroitin 6-sulphotransferase (CGST) and using this as a probe in
 CC a random oligonucleotide-primed labelling method. Human foetal brain
 CC cDNA was inserted into a phage lambda g11 cloning vector whereby the
 CC clones containing the KSGal6ST were obtained by hybridization using the
 CC prepared probe. The positive clones were subcloned into a recombinant
 CC expression vector and used to transform COS-7 cells, from which cells,
 CC expressing KSGal6ST can be selected. The phage cDNA inserts were
 CC isolated and subcloned into a Bluescript plasmid. Deletion clones were
 CC then prepared from which both strands were sequenced by the Sanger
 CC method. The KSGal6ST of the invention transfers the sulphate from a
 CC sulphate donor to galactose 6-OH groups in keratan sulphate, but does not
 CC transfer sulphate to chondroitin, chondroitin sulphate A or C, dermatan
 CC sulphate or CDSNS heparin.
 XX
 SQ Sequence 411 AA;
 XX
 Query Match 100.0%; Score 31; DB 19; Length 411;
 Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 300 VRYEDL 305
 XX
 RESULT 12
 ID ABB81557 standard; Protein; 418 AA.
 XX
 AC ABB81557;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Mouse intestinal N-acetylglucosamine-6-sulfoltransferase SEQ ID NO:5.
 XX
 KM Human; N-acetylglucosamine-6-sulfoltransferase; enzyme; GlcNAc6ST;
 KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 XX

KM ophthalmological.
 XX
 OS Mus musculus.
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 09-AUG-2001; 2001US-0927602.
 XX
 PR 11-AUG-2000; 2000US-325773P.
 XX
 PA (FUKU/) FUKUDA M.N.
 PA (AKAM/) AKAMA T.O.
 XX
 PI Fukuda MN, Akama TO;
 XX
 DR WPI; 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfoltransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX
 PS Example 5; Page 24-25; 69pp; English.
 XX
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfoltransferase (I), which is able to catalyse
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents mouse intestinal N-acetylglucosamine-6-
 CC sulfoltransferase, which is given in comparison with (I) in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 418 AA;
 XX
 Query Match 100.0%; Score 31; DB 23; Length 418;
 Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 294 VRYEDL 299
 XX
 RESULT 13
 ID AAY31656 standard; Protein; 483 AA.
 XX
 AC AAY31656;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Mouse N-acetylglucosamine-6-O-sulfoltransferase.
 XX
 KM N-acetylglucosamine-6-O-sulfoltransferase; mouse; GlyCAM-1;
 KM L-selectin ligand.
 XX
 OS Mus musculus.
 XX
 PN EP943688-A2.
 XX
 PD 22-SEP-1999.
 XX
 PF 04-MAR-1999; 99BP-0301530.
 XX
 PR 24-JUN-1998; 98JP-0177844.
 PR 05-MAR-1996; 98JP-0054007.
 XX
 PA (SECK) SEIKAGAKU CORP.
 XX

```

XX Habuchi O, Kadomatsu K, Kannagi R, Muramatsu H;
PI Muramatsu T, Uchimura K;
XX WPI: 1999-520337/44.
DR N-PSDB: AAX87820.
XX
PT New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PT for synthesis of sugar chains, e.g. GLYCAM-1
XX
PS Claim 1; Page 24-25; 41pp; English.
XX
CC The present sequence represents mouse N-acetylglucosamine-6-O-
CC sulfofructanase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC beta 1-3 = beta 1-3 glycosidic linkage, and beta 1-4 = beta 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GlyCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87820) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GlyCAM-1 using transformants which harbour
CC the DNA.
XX
SQ Sequence 483 AA;
XX
Query Match 100.0%; Score 31; DB 20; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VREYDL 6
DB 376 VREYDL 381
XX
RESULT 14
AAX31657
ID AAX31657 standard; Protein; 484 AA.
XX
AC AAX31657;
XX
DT 09-NOV-1999 (first entry)
XX
DE Human N-acetylglucosamine-6-O-sulfotransferase.
XX
KM N-acetylglucosamine-6-O-sulfotransferase; human; GlyCAM-1;
XX L-selectin ligand.
XX
OS Homo sapiens.
XX
PN EP943668-A2.
XX
PD 22-SEP-1999.
XX
PF 04-MAR-1999; 99EP-0301530.
XX
PR 24-JUN-1998; 98JP-0177844.
XX
PR 05-MAR-1998; 98JP-0054007.
XX
PA (SEKK) SEIKAGAKU CORP.
XX
PI Habuchi O, Kadomatsu K, Kannagi R, Muramatsu H;
XX Muramatsu T, Uchimura K;
XX WPI: 1999-520337/44.
XX DR N-PSDB: AAX87821.
XX
PT New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PT for synthesis of sugar chains, e.g. GlyCAM-1

```

```

XX Claim 2; Page 28-30; 41pp; English.
XX
CC The present sequence represents human N-acetylglucosamine-6-O-
CC sulfofructanase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC beta 1-3 = beta 1-3 glycosidic linkage, and beta 1-4 = beta 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GlyCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87821) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GlyCAM-1 using transformants which harbour
CC the DNA.
XX
SQ Sequence 484 AA;
XX
Query Match 100.0%; Score 31; DB 20; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VREYDL 6
DB 377 VREYDL 382
XX
RESULT 15
AAB95367
ID AAB95367 standard; Protein; 530 AA.
XX
AC AAB95367;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17679.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 17679; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

```

CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX SQ Sequence 530 AA;

XX Query Match 100.0%; Score 31; DB 22; Length 530;

XX Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 VRYEDL 6
XX |||||

Db 423 VRYEDL 428

RESULT 16

AAU69414 standard; Protein; 531 AA.

XX AC AAU69414;

XX DT 30-JAN-2002 (first entry)

XX DE Lung small cell carcinoma antigen #8.

XX KM Human; cytostatic; antitumour; lung small cell cancer antigen;

XX KW tumour; lung cancer.

XX OS Homo sapiens.

XX PN WO200177168-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-US11859.

XX PR 11-APR-2000; 2000US-196780P.

XX PR 21-JUN-2000; 2000US-213361P.

XX PR 01-SEP-2000; 2000US-229763P.

XX PR 05-SEP-2000; 2000US-230629P.

XX PR 14-SEP-2000; 2000US-232565P.

XX PR 19-DEC-2000; 2000US-257037P.

XX PR 08-JAN-2001; 2001US-260796P.

XX PA (CORI-) CORIXA CORP.

XX PI Lodes M.J., Wang T., Mohamath R., Indirias C.Y.

XX DR WPI; 2002-010896/01.

XX PT Lung tumour polynucleotide and polypeptides useful in therapy and
XX diagnosis of cancer especially lung cancer
XX Claim 1; Page 219-220; 295pp; English.
XX The invention relates to novel isolated lung small cell cancer antigen
XX polynucleotides (I) and polypeptides (II) used in a method of detecting

CC cancer in a patient. The method is optionally performed by
CC utilising oligonucleotides (III), where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAU69407-AAU69431 represent novel human lung small cell
CC cancer antigen amino acid sequences of the invention.

XX SQ Sequence 531 AA;

XX Query Match 100.0%; Score 31; DB 23; Length 531;

XX Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 VRYEDL 6
XX |||||

Db 424 VRYEDL 429

RESULT 17

ABB67037 standard; Protein; 649 AA.

XX AC ABB67037;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 27903.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter J.C., Adams M., Li P.M.D., Myers E.W.

XX PR WPI: 2001-656660/75.

XX PR N-PSDB; ABL11140.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Disclosures: SEQ ID NO 27903; 21pp + Sequence Listing; English.XX PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
XX
XX
SQ Sequence 649 AA;

Query Match 100.0%; Score 31; DB 22; Length 649;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

DB 332 IRYEDL 337

RESULT 18

AB64512
ID ABB64512 standard; Protein: 363 AA.

XX
XX ABB64512;

XX
XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 20328.

XX
XX Drosophila; developmental biology; cell signalling; insecticide;

XX
XX pharmaceutical.

OS Drosophila melanogaster.

XX
XX WO200171042-A2.

XX
XX 27-SEP-2001.

XX
XX 23-MAR-2001; 2001WO-US09231.

XX
XX 23-MAR-2000; 2000US-19137P.

XX
XX 11-JUL-2000; 2000US-0614150.

XX
XX (PEKE) PE CORP NY.

XX
XX Venter JC, Adams M, Li PWD, Myers EW;

XX
XX WPI; 2001-656860/75.

XX
XX N-PSDB; ABL08615.

XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX
XX Disclosure; SEQ ID NO 20328; 21pp + Sequence Listing; English.

XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of

XX
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins

XX
XX (AB57737-AB872072).

XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX
XX SQ Sequence 363 AA;

Query Match 96.8%; Score 30; DB 22; Length 363;

Best Local Similarity 83.3%; Pred. No. 1.9e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

DB 258 IRYEDL 263

RESULT 19

AB609452
ID ABB09452 standard; Protein: 562 AA.

XX
XX ABB09452;

XX
XX 01-JUL-2002 (first entry)

DE Glucose-3-dehydrogenase amino acid sequence.

XX
XX Glucose-3-dehydrogenase; G3DH; enzyme.

XX
XX Halomonas sp. alpha-15.

XX
XX JP2002017372-A.

XX
XX 22-JAN-2002.

XX
XX 30-JUN-2000; 2000JP-0237709.

XX
XX 30-JUN-2000; 2000JP-0237709.

XX
XX (HAYADA) HAYADA K.

XX
XX WPI; 2002-221711/28.

XX
XX N-PSDB; ABL52744.

XX
XX Glucose-3-dehydrogenase and its preparation -

XX
XX Claim 1a; Page 7-9; 16pp; Japanese.

XX
XX The invention relates to Glucose-3-dehydrogenase (G3DH) enzyme. The
XX methods of the invention are used for the preparation of G3DH. The
XX current sequence represents the glucose-3-dehydrogenase amino acid
XX sequence.

XX
XX SQ Sequence 562 AA;

Query Match 96.8%; Score 30; DB 23; Length 562;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

DB 142 IRYEDL 147

RESULT 20

AB692113
ID ABB92113 standard; Protein: 405 AA.

XX
XX ABB92113;

XX
XX 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1324.

XX
XX Herbicidally active polypeptide; herbicide.

XX
XX Arabidopsis thaliana.

XX
XX WO200210210-A2.

XX
XX 07-FEB-2002.

XX
XX 28-AUG-2001; 2001WO-EP09892.

XX
XX 28-AUG-2001; 2001WO-EP09892.

XX
XX (FARB) BAYER AG.

PI Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 PS Claim 5; SEQ ID NO 1324; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins
 CC (AB930790-AB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 CC
 SQ Sequence 405 AA;
 XX
 QY Query Match 93.5%; Score 29; DB 23; Length 405;
 Best Local Similarity 83.3%; Pred. No. 3.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 335 VREEDL 6
 1 VREEDL 6
 |||||
 ABP32235
 ID ABP32235 standard; Protein; 139 AA.
 XX
 AC ABP32235;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human isomerase-like ORF1208 protein, SEQ ID NO:2416.
 XX
 KW Human: ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antiproliferative; antidiabetic; cyostatic; nootropic;
 KW neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shinkets RA;
 XX
 DR WPI; 2002-106200/14.
 DR N-PDB; ABN76261.
 XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 10; Page 866; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antinflammatory activity, tumour inhibition activity,
 CC and antinecrotic activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 SQ Sequence 139 AA;
 XX
 QY Query Match 90.3%; Score 28; DB 23; Length 139;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 123 VREEDL 6
 1 VREEDL 6
 |||||
 ABB81563
 ID ABB81563 standard; Protein; 174 AA.
 XX
 AC ABB81563;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Human chondroitin-6-sulfotransferase portion SEQ ID NO:11.
 XX
 KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme; GICNAC6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX

PF 09-AUG-2001; 2001US-0927602.
 XX
 PR 11-AUG-2000; 2000US-325773P.
 XX
 PA (FUKU/) FUKUDA M.N.
 PA (AKAM/) AKAMA T.O.
 XX
 PI Fukuda MN, Akama TO;
 XX
 DR WPI; 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX
 PS Example 2; Fig 3C; 69pp; English.
 XX
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratotomy. The present
 CC sequence represents a portion of the human chondroitin-6-
 CC sulfotransferase, which is given in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 174 AA;
 XX
 Query Match 90.3%; Score 28; DB 23; Length 174;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 Db 132 VRYEDV 137
 XX
 RESULT 23
 ABB55144
 ID ABB55144 standard; Protein; 233 AA.
 XX
 AC ABB55144;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein ysfD.
 XX
 KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bojoline A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR WPI; 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species -
 XX
 PS Claim 6; SEQ ID NO 1846; 2504pp; French.
 XX

CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO20017734 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 233 AA;
 XX
 Query Match 90.3%; Score 28; DB 23; Length 233;
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 Db 146 IRYEDI 151
 XX
 RESULT 24
 ABB53027
 ID ABB53027 standard; Protein; 254 AA.
 XX
 AC ABB53027;
 XX
 DT 11-FEB-2002 (first entry)
 XX
 DE Escherichia coli polypeptide SEQ ID NO 1422.
 XX
 KM Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KM immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicaemia;
 KW pyelonephritis; antibiotic resistance.
 XX
 OS Escherichia coli.
 XX
 PN WO200166572-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-EP03445.
 XX
 PR 10-MAR-2000; 2000FR-0003145.
 PR 02-FEB-2001; 2001FR-0001449.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
 XX
 DR WPI; 2001-550253/61.
 XX
 PT A library of DNA fragments of Escherichia coli strains for the
 PT phylogenetic determination of a given strain comprises polynucleotides of
 PT nature B2/D+ A- -
 XX
 PS Example 6; Fig 6; 645pp; English.
 XX
 CC The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA899533)
 CC and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli
 CC infections. The polypeptides are useful for determining the phylogenetic
 CC group of a given E. coli strain. These polypeptides can detect and treat
 CC an undesired development of E. coli, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicaemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more

CC frequent use of broad spectrum antibiotics.
 XX
 SQ Sequence 254 AA;

Query Match 90.3%; Score 28; DB 22; Length 254;
 Best Local Similarity 83.3%; Pred. No. 3.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Caps 0;

OY 1 VRYEDL 6
 |||||
 Db 143 VRYQDL 148

RESULT 25
 AAR83040
 ID AAR83040 standard; peptide; 306 AA.
 XX
 AC AAR83040;
 XX

DT 03-APR-1996 (first entry)
 XX
 XX

DE Capsular polysaccharide of Streptococcus pneumoniae (cpsu).
 XX
 KM Capsular polysaccharide; cps; peptide; flanking region; detection;
 KM serotype; diagnosis; prevention; Streptococcus pneumoniae.
 XX

OS Streptococcus pneumoniae.
 XX

PN W09531548-A1.
 XX

PD 23-NOV-1995.
 XX

PF 16-MAY-1995; 95MO-US06119.
 XX

PR 16-MAY-1994; 94US-0243546.
 XX

PA (UABR-) UAB RES FOUND.
 XX

PI Dillard J, Yother J;
 XX

DR WPI: 1996-010934/01.
 XX

DR N-PSDB; AAT05848.
 XX

PT New Streptococcus pneumoniae capsular polysaccharide genes - used
 for detection, serotyping and for diagnosis and prevention of S.
 pneumoniae infection
 XX

PS Disclosure; Page 178-180; 226pp; English.
 XX

CC Sequences encoding the 5' flanking region of the capsular
 CC polysaccharide gene (cps) of Streptococcus pneumoniae and which are
 CC of sufficient length to allow hybridisation under standard
 CC hybridisation conditions to a S.pneumoniae cps gene flanking region
 CC may be used in methods to detect and serotype S.pneumoniae. They
 CC may also be used for the diagnosis and prevention of S. pneumoniae
 CC infection.
 CC

SQ Sequence 306 AA;

Query Match 90.3%; Score 28; DB 17; Length 306;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 165 VRYEDV 170

RESULT 26
 AAY33895
 ID AAY33895 standard; Protein; 306 AA.
 XX
 AC AAY33895;

XX 04-JAN-2000 (first entry)
 DT
 XX
 DE Amino acid sequence of Cps3U.
 XX

KW Infection; bacterial capsule; serotype specific; virulence;
 KW diagnosis; capsule synthesis; cassette model; binary model;
 KW cassette-in; flanking region; non-serotype specific.
 XX

OS Streptococcus pneumoniae.
 XX

PN US5948900-A.
 XX

PD 07-SEP-1999.
 XX

PF 02-JUN-1997; 97US-0867030.
 XX

PR 02-JUN-1997; 97US-0867030.
 XX

PR 16-MAY-1994; 94US-0243546.
 XX

PA (UABR-) UAB RES FOUND.
 XX

PI Dillard J, Yother J;
 XX

DR WPI: 1999-517979/43.
 XX

DR N-PSDB; AA211802.
 XX

PT Non-serotype-specific flanking nucleotide sequences derived from
 Streptococcus pneumoniae capsular polysaccharide genes, useful as
 PT hybridization probes for identifying serotype-specific capsular
 PT polysaccharide genes
 XX

PS Example 10; Fig 6g; 140pp; English.
 XX

CC This is the amino acid sequence of the capsule synthesis protein,
 CC Cps3U.
 CC

CC Serotype-specific cps genes encode the various enzymatic
 CC functions of capsule synthesis and hence determine the particular
 CC structure of the capsule polysaccharides produced. Therefore cps genes
 CC determine the serotype of the virus.
 CC

CC The non-serotype specific flanking nucleotide sequences (AA211798,
 CC AA211799, AA211800, AA211801 and AA211802) may be used in hybridization
 CC assays to identify the location of DNA flanking serotype-specific cps
 CC genes in any strain of S. pneumoniae virus.
 CC

CC Additionally, the flanking regions are involved in recombination
 CC and integration of the type specific cps genes during virus
 CC replication. Therefore, when a selected gene (such as the
 CC serotype-specific cps genes) is integrated between the nucleic acids
 CC of the flanking region, the resulting construct can be stably integrated
 CC into an S.pneumoniae host.
 CC

CC This allows the locus of the serotype-specific cps genes (and other
 CC genetic elements), within a virus, to be identified, and therefore,
 CC isolated and characterized. The cps gene sequences may then be used as
 CC hybridization probes to identify and quantify specific serotypes of S.
 CC pneumoniae or to recombinantly produce the gene products for use as
 CC antigens in immunization protocols.
 CC

SQ Sequence 306 AA;

Query Match 90.3%; Score 28; DB 20; Length 306;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 165 VRYEDV 170

RESULT 27
 ABB91753
 ID ABB91753 standard; Protein; 333 AA.
 XX
 AC ABB91753;

XX 31-MAY-2002 (first entry)
DT
XX
DE Herbicidally active polypeptide SEQ ID NO 964.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PE 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI, 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 964; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (AB990790-AB994016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 333 AA:
Query Match 90.3%; Score 28; DB 23; Length 333;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VREEDL 6
Db 232 MREEDL 237
RESULT 28
AAG44207
ID AAG44207 standard; Protein: 349 AA.
XX
AC AAG44207;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55345.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130831.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154038.
PR 16-SEP-1999; 99US-0154339.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 349;
Best Local Similarity 93.3%; Pred. No. 5e+02; 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 1 VRYEDL 6
:|||||
Db 244 LRYEDL 249

RESULT 29
AAC46838
ID AAC46838 standard; Protein; 349 AA.
XX
XX AAC46838;
AC
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 58967.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150984.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 349;
 Best Local Similarity 83.3%; Pred. No. 56+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 244 LRYEDL 249

RESULT 30
 AAG44206
 ID AAG44206 standard; Protein; 354 AA.
 XX

AC AAG44206;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55344.

KW Protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

OS Arabidopsis thaliana.

XX
 PN EP1033405-A2.

XX
 PD 06-SEP-2000.

XX
 PE 25-FEB-2000; 2000EP-0301439.

XX
 PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 14-MAY-1999; 99US-0134768.
 PR 18-MAY-1999; 99US-0134941.
 PR 19-MAY-1999; 99US-0135124.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144864.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159289.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159639.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 354;
Best Local Similarity 83.3%; Pred. No. 5e+02; Mismatches 0; Gaps 0;
Matches 5; Conservative 1; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 31
AAG46837
ID AAG46837 standard; Protein; 354 AA.
XX
AC AAG46837;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58966.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134770.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140391.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158365.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 354;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 32
AB93382
ID AB93382 standard; Protein: 354 AA.

AC AB93382;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2593.

KW Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -

PS Claim 5; SEQ ID NO 2593; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
CC (AB930790-AB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

XX Sequence 354 AA;

Query Match 90.3%; Score 28; DB 23; Length 354;

Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 33
AAG44205
ID AAG44205 standard; Protein: 359 AA.

AC AAG44205;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55343.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

```
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149472.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 359;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDDL 6
DB 254 LREDDL 259

RESULT 34
```


PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159631.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 359;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 254 LRYEDL 259

RESULT 35
AAB49722
ID AAB49722 standard; Protein; 359 AA.
XX
AC AAB49722;
XX
DT 09-APR-2001 (first entry)
XX

DE Hydroxyjasmonic acid sulfotransferase AtST2a protein.
XX
KW Hydroxyjasmonic acid sulfotransferase; AtST2a; flowering time; cabbage;
KW jasmonate; genetically modified plant; lettuce; sugar cane; carrot;
KW increase vegetative growth; biomass increase.
XX
OS Arabidopsis thaliana.
XX
PN W0200102589-A2.
XX
PD 11-JAN-2001.
XX
PF 06-JUL-2000; 200GMO-CA08001.
XX
PR 06-JUL-1999; 99CA-2274873.
XX
PA (VARI/) VARIN L.
XX (GIDD/) GIDDA S.
PI Varin L, Gidda S;
XX
DR WPI: 2001-159272/16.
DR N-PSDB; AAF29177.
XX
PT Methods for modulating flowering in plants, particularly useful for
PT plants used in the food-processing industry, involves modifying the
PT endogenous level of compounds of the jasmonate family -
XX
PS Claim 50; Fig 8; 50pp; English.
XX
CC This invention relates to a method for modulating flowering in a plant.
CC The method comprises modifying the endogenous level of at least one
CC compound of the jasmonate family in a plant. The methods are used to
CC produce plants which are genetically modified to flower early or tardily
CC when compared to a corresponding plant that is not genetically modified,
CC where the modified plant has an increased (flower early) or lowered
CC (flower tardily) level of jasmonic acid, or a compound of the jasmonate
CC family. The method is useful for modulating flowering, particularly for
CC plants that are used in the food-processing industry and plants with
CC horticultural value. The method is particularly useful for e.g. delaying
CC flowering time in crops like lettuce, cabbage, sugar cane or carrots.
CC which results in increased vegetative growth and biomass. The present
CC sequence represents the Arabidopsis thaliana AtST2a protein which is an
CC hydroxyjasmonic acid sulfotransferase. The AtST2a gene can be used in
CC the method of the invention.
CC
SQ Sequence 359 AA;
XX

Query Match 90.3%; Score 28; DB 22; Length 359;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 254 LRYEDL 259

RESULT 36
AA939919
ID AA939919 standard; Protein; 388 AA.
XX
AC AA939919;
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 protein sequence.
XX
KW Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KW selectin binding interaction; inflammation; lymphocyte homing; mouse;
KW secondary lymph organ.
XX
OS Mus sp.
XX


```

PN FR2792651-A1.
PD 27-OCT-2000.
PE 21-APR-1999; 99FR-0005034.
PF 21-APR-1999; 99FR-0005034.
PR 21-APR-1999; 99FR-0005034.
PS (CNRS ) CNRS CENT NAT RECH SCI
PT (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Hellig R;
XX WPI, 2001-126236/14.
XX
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
XX proteins useful in industry -
XX
XX Claim 7; Pages 1431-1433; 1657pp; French.
XX
XX The present invention relates to the genomic sequence of Pyrococcus
XX abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
XX a hyperthermophilic archaeon, which is isolated from deep-sea
XX hydrothermal vents. The present sequence is one such P. abyssi protein.
XX The proteins of the present invention have various potential industrial
XX uses, since the proteins are stable at very high temperatures, some up to
XX 110 degrees centigrade.
XX Note: This patent is in the same patent family as WO200065062, which
XX contains additional sequences as shown in AAB99132-AAB99143,
XX AAB73903-AAB75920 and AAG66436.
XX
XX Sequence 458 AA;
XX
XX Query Match 90.3%; Score 28; DB 22; Length 458;
XX Best Local Similarity 83.3%; Pred. No. 6.7e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 VRYEDL 6
XX :|||||
XX DB 25 LRYEDL 30
XX
XX RESULT 39
XX AAM52863
XX ID AAM52863 standard; Protein; 479 AA.
XX
XX AC AAM52863;
XX
XX 30-JUN-1998 (first entry)
XX
XX Glycosaminoglycan sulphotransferase protein.
XX
XX Glycosaminoglycan sulphotransferase; C65T; N-acetylglucosamine;
XX galactose; glycosaminoglycan; chondroitin; keratan sulphate; diagnostic.
XX
XX Homo sapiens.
XX
XX EP821066-A1.
XX
XX 28-JAN-1998.
XX
XX 22-JUL-1997; 97EP-0305476.
XX
XX 24-JUL-1996; 96JP-0195063.
XX
XX (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
XX Fukuta M, Habuchi O;
XX
XX WPI, 1998-102622/10.
XX
XX Human glycosaminoglycan sulpho:transferase protein - catalyses

```

PT	6-sulphation of chondroitin to produce chondroitin sulphate for use
PT	in pharmaceutical(s)
XX	
PS	Claim 3; Page 21-23; 27p; English.
XX	
CC	The human glycosaminoglycan sulphotransferase (C6ST) is a type II
CC	membrane protein derived from foetal brain tissue with a molecular weight
CC	of 50-55 kD. The protein transfers sulphate groups from a donor to the
CC	N-acetylglucosamine residue or galactose residue of glycosaminoglycan.
CC	C6ST displays substrate specificity transferring the sulphate group to
CC	the hydroxyl group position at C-6 of the N-acetylglucosamine residue
CC	of chondroitin and the hydroxyl group position at C-6 of the galactose
CC	residue of keratan sulphate. The protein is useful for studying the
CC	function of chondroitin sulphate and can provide chondroitin sulphate
CC	that may be useful in pharmaceuticals. Both the protein and DNA might
CC	be useful for treating or diagnosing diseases attributable to low C-6
CC	sulphation of chondroitin N-acetylglucosamine residues.
XX	
SQ	Sequence 479 AA;
XX	
Query Match	90.3%; Score 28; DB 19; Length 479;
Best Local Similarity	83.3%; Pred. No. 7.1e+02;
Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 VRYEDL 6
DB	369 VRYEDV 374
XX	
RESULT 40	
AA66031	
ID	AA66031 standard; Protein: 567 AA.
XX	
AA66031;	
XX	
28-JUN-1995	(first entry)
XX	
DE	Mammalian MEK kinase (MEK 3).
XX	
MEK kinase; MEK 3; mitogen-activated protein kinase regulator;	
XX	MAPK; cell atrophy inhibition; Parkinson's; Alzheimer's; cancer;
KW	autoimmune diseases; allergies; wound healing; oncogenes;
KW	tumour agents; neurotropic growth factor.
XX	
OS	Mus musculus.
XX	
Key	Location/Qualifiers
XX	
FT	Domain
XX	
FT	1..174
XX	/note="amino terminus regulatory domain"
FT	175..179
XX	/note="regulatory hinge region"
FT	357..567
XX	/note="catalytic domain"
XX	
XX	WO9424159-A.
XX	
27-OCT-1994.	
XX	
15-APR-1994;	94WO-US04178.
XX	
15-APR-1993;	93US-0049254.
XX	
(NAME-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.	
XX	
Johnson GL;	
XX	
WPI; 1994-357747/44.	
XX	
New MEK kinase protein and related antibodies and nucleic acid	
XX	regulator of mitogen activated protein kinase, useful
XX	therapeutically to inhibit cell atrophy, to screen for oncogenes
XX	etc.
XX	

PS Claim 6; Page 13; 84pp; English.

XX AA079325 encodes AAR66029 the mammalian MEK kinase (MEK 1), other
 CC unique mammalian MEK kinases identified by PCR are described in
 CC AAR66030 (MEK 2), AAR66031 (MEK 3) and AAR66032 (MEK 4). MEK is an
 CC activator independent of Raf protein, of mitogen-activated protein
 CC kinases (MAPK). Inactivation of MEK can be used in the treatment
 CC of some cancers, autoimmune diseases and allergies, while
 CC stimulation can promote wound healing. MEK can also be used to
 CC alleviate cellular atrophy in Parkinson's or Alzheimer's by acting
 CC as a neurotrophic growth factor, and to screen for oncogenes and
 CC tumour agents.

SQ Sequence 567 AA;

Query Match 90.3%; Score 28; DB 15; Length 567;
 Best Local Similarity 83.3%; Pred. No. 8.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 64 VRYEDV 69

RESULT 41
 ABB92573
 ID ABB92573 standard; Protein: 612 AA.

XX AC ABB92573;
 XX 31-MAY-2002 (first entry)
 XX DE Herbicidally active polypeptide SEQ ID NO 1784.
 XX KW Herbicidal: plant; agriculture; herbicide.
 XX OS Arabidopsis thaliana.
 XX PN WO200210210-A2.
 XX PD 07-FEB-2002.
 XX PF 28-AUG-2001; 2001WO-EP09892.
 XX PR 28-AUG-2001; 2001WO-EP09892.
 XX PA (FARB) BAYER AG.
 XX PT Tietjen K, Weidner M;
 XX PI Tietjen K, Weidner M;
 XX DR WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds,
 XX comprising aligning and comparing nucleic acid or amino acid sequences
 XX from plant with nucleic acid or amino acid sequences from non-plant
 XX organisms -
 XX PS Claim 5; SEQ ID NO 1784; 261pp + Sequence Listing; English.

XX CC The invention relates to identifying target proteins
 XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
 XX aligning and comparing nucleic acid or amino acid sequences from plant
 XX with nucleic acid or amino acid sequences from non-plant organisms using
 XX suitable search parameters, where plant sequences having an E-value
 XX greater by a factor of 3 than the E-value of most similar non-plant
 XX sequences are selected. The polypeptides or nucleic acids encoding them
 XX are useful for identifying modulators. The identified modulators are
 XX useful as herbicides.

SQ Sequence 612 AA;

Query Match 90.3%; Score 28; DB 23; Length 612;
 Best Local Similarity 83.3%; Pred. No. 9.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 :|||||
 Db 230 LRYEDL 235

RESULT 42
 AAR77546
 ID AAR77546 standard; Protein: 626 AA.

XX AC AAR77546;
 XX 12-MAR-1996 (first entry)
 XX DE MEK3 protein.
 XX KW MEK3; mitogen ERK kinase kinase; signal transduction; homeostasis;
 XX cancer; inflammation; neurological disorder; autoimmune disease;
 XX allergy; hormone-related disease.
 XX OS Mus sp.
 XX Key Location/Qualifiers
 XX FH 1..174
 XX FT Domain /label= Regulatory_domain
 XX FT Region 175..179
 XX FT /label= Regulatory_hinge_sequence
 XX FT Domain 357..626
 XX /label= Catalytic_domain

XX PN WO9528421-A1
 XX PD 26-OCT-1995.
 XX PF 14-OCT-1994; 94WO-US11690.
 XX PR 15-APR-1994; 94WO-US04178.
 XX PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 XX PI Johnson GL;
 XX DR WPI; 1995-373762/48.
 XX DR N-PSDB; AAT05572.

XX PT New signal-regulated kinase proteins and nucleic acids - used for
 XX regulating cell responsiveness for treating e.g. tumours,
 XX auto-immune disease, inflammation or neuronal disorders.
 XX PS Claim 1; Page 31-34; 147pp; English.

XX CC MEK2, MEK3, MEK4 and MEK5 (AAR77545-48) are mitogen ERK kinase
 XX kinases capable of regulating the activity of mitogen-activated protein
 XX kinase. The proteins can be produced in heterologous host cells by
 XX expression of cDNAs (AAT05571-74) isolated from PC12 and HL60 cells.
 XX CC They are capable of regulating signal transduction in cells. By
 XX modulating the activity of an MEK-dependent pathway relative to that of
 XX the Raf-dependent pathway it is possible to inhibit cell growth for
 XX treatment of cancer, autoimmunity, etc.

SQ Sequence 626 AA;

Query Match 90.3%; Score 28; DB 16; Length 626;
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 64 VRYEDV 69

RESULT 43

AAW56159
ID AAW56159 standard; Protein; 626 AA.
XX
AC AAW56159;
XX
DT 17-JUL-1998 (first entry)
XX
DE A mitogen-activated protein kinase kinase (MAPKK).
XX
KW Mitogen-activated protein kinase kinase; MAPKK; mouse;
KW extracellular signal-regulated kinase kinase; MEKK; regulation;
KW signal transduction; raf-independent arm; screening assay; treatment;
KW disorder; cancer; autoimmune disease; inflammation; allergy;
KW neuronal disease; Parkinson's disease; Alzheimer's disease; ds.
XX
OS Mammalia.
XX
PN US5753446-A.
XX
PD 19-MAY-1998.
XX
PF 06-JUN-1995; 95US-0472934.
XX
PR 15-APR-1993; 93US-0049254.
PR 14-OCT-1994; 94US-0323460.
PR 21-FEB-1995; 95US-0354516.
PR 12-MAY-1995; 95US-0440421.
XX
PA (NAE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Johnson GL;
XX
PT WPI; 1998-311395/27.
DR N-PSDB; AAW22678.
XX
PT Screening assay for regulators of MEKK signal transduction - using
PT mammalian MEKK polypeptide
XX
PS Claim 13; Columns 51-54; 48pp; English.
XX
CC The present sequence represents a mitogen-activated protein kinase
CC kinase (MAPKK) (also known as extracellular signal-regulated kinase
CC kinase (MEKK)). The protein, which is serine/threonine kinase is capable
CC of regulating signal transduction in cells. It regulates the activity of
CC elements of the raf-independent arm of MEKK. A screening assay for
CC compounds that regulate signal transduction by a MEKK protein comprises
CC contacting a reaction mixture containing a mammalian MEKK polypeptide and
CC a test compound and determining the effect of the test compound on an
CC indicator of signal transduction by the MEKK polypeptide in the reaction
CC mixture. Compounds identified by the above assay can be used to prepare
CC therapeutic compositions for treating disorders that are subject to
CC regulation or cure by manipulating a signal transduction pathway in
CC cells involved in the disorders, e.g. cancer, autoimmune diseases,
CC inflammations, allergies, and neuronal diseases such as Parkinson's
CC disease and Alzheimer's disease.
XX
SQ Sequence 626 AA;
XX
Query Match 90.3%; Score 28; DB 19; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRYEDL 6
DB 64 VRYEDV 69
XX
RESULT 44
AAV43320
ID AAV43320 standard; Protein; 626 AA.
XX
AC AAV43320;
XX

DT 24-JAN-2000 (first entry)
XX
DE Mitogen ERK kinase kinase, MEKK-3, protein sequence.
XX
KW Mitogen ERK kinase kinase; MEKK; MEKK-3; neurological disorder; cancer;
KW extracellular signal-regulated kinase; inflammation; autoimmune disease;
KW allergic reaction; hormone related disease; therapy.
XX
OS Homo sapiens.
XX
PN US5981265-A.
XX
PD 09-NOV-1999.
XX
PF 05-JUN-1995; 95US-0461146.
XX
PR 15-APR-1993; 93US-0049254.
PR 12-MAY-1995; 95US-0440421.
PR 15-APR-1994; 94WO-US04178.
PR 14-OCT-1994; 94US-0323460.
PR 14-OCT-1994; 94WO-US11690.
PR 28-NOV-1994; 94US-0345516.
XX
PA (NAE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Johnson GL;
XX
PT WPI; 1999-633328/54.
DR N-PSDB; AAZ31879.
XX
PT Regulating mitogen extracellular signal-regulated kinase kinase protein
PT activity, useful for the treatment of cancer, neurological diseases and
PT autoimmune diseases -
XX
PS Claim 2; Column 65-70; 94pp; English.
XX
CC This sequence is the mitogen ERK (extracellular signal-regulated
CC kinase) kinase kinase-3 (MEKK-3). The invention relates to a method of
CC regulating MEKK protein activity in a fungal cell by transforming or
CC transfecting the cell with a nucleic acid encoding an MEKK protein. The
CC MEKK protein is useful for treating cancer, inflammation, neurological
CC disorders, autoimmune diseases, allergic reactions, and hormone related
CC diseases.
XX
SQ Sequence 626 AA;
XX
Query Match 90.3%; Score 28; DB 20; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRYEDL 6
DB 64 VRYEDV 69
XX
RESULT 45
AAV42106
ID AAV42106 standard; Protein; 626 AA.
XX
AC AAV42106;
XX
DT 09-DEC-1999 (first entry)
XX
DE Human MEKK3 protein sequence.
XX
KW MEKK1; MEKK2; MEKK3; mitogen-activated protein kinase; MAPK; ERK;
KW extracellular regulated kinase; signal transduction; regulation;
KW MAPK/ERK; MEK; MEKK; inflammation; cellular proliferation;
KW differentiation; development; cell death.
XX
OS Homo sapiens.
XX
PN WO9947686-A2.
XX

```

XX 23-SEP-1999.
PD 15-MAR-1999; 99WO-US05556.
XX 16-MAR-1998; 98US-0078153.
PR 04-SEP-1998; 98US-0099165.
XX (CADU-) CADUS PHARM CORP.
PA Johnson GL;
PI WPI: 1999-571843/48.
DR N-PSDB; AA225071.
XX New human MEK polynucleotides and polypeptides, used for regulating
PT signal transduction in cells -
PS Claim 12; Fig 10; 159pp; English.
XX The present sequence represents human mitogen-activated protein kinase/
CC extracellular response kinase (MAPK/ERK) kinase kinase (MEKK),
CC specifically designated MEKK3. The MEKK proteins are used to modulate
CC and regulate signal transduction in cells, as well as for regulation of
CC gene transcription in a cell encoding MEKK, where the cell is involved
CC in inflammation, regulation of cellular proliferation and
CC differentiation, regulation of development, regulation of cell death or
CC regulation of inflammation. They are also used to prepare antibodies.
CC MEKK polynucleotides can be used to produce the protein recombinantly
CC and as a source of probes and primers.
XX
SQ Sequence 626 AA;
Query Match 90.3%; Score 28; DB 20; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRYEDL 6
Db 64 VRYEDV 69
RESULT 46
AAV42109
ID AAV42109 standard; Protein: 626 AA.
XX AAV42109;
AC AAV42109;
XX 09-DEC-1999 (first entry)
DT Murine MEKK3 protein sequence.
DE MEKK1; MEKK2; MEKK3; mitogen-activated protein kinase; MAPK; ERK;
XX extracellular response kinase; signal transduction; regulation;
XX MAPK/ERK; MEK; MEKK; inflammation; cellular proliferation;
KW differentiation; development; cell death.
XX Mus musculus.
OS
XX WO9947686-A2.
PN 23-SEP-1999.
XX 15-MAR-1999; 99WO-US05556.
XX 16-MAR-1998; 98US-0078153.
PR 04-SEP-1998; 98US-0099165.
XX (CADU-) CADUS PHARM CORP.
PA Johnson GL;
PI WPI: 1999-571843/48.
DR

```

```

DR N-PSDB; AA225074.
XX New human MEK polynucleotides and polypeptides, used for regulating
PT signal transduction in cells -
XX Example 3; Fig 12; 159pp; English.
XX The present sequence represents murine mitogen-activated protein kinase/
CC extracellular response kinase (MAPK/ERK) kinase kinase (MEKK),
CC specifically designated MEKK3. The MEKK proteins are used to modulate
CC and regulate signal transduction in cells, as well as for regulation of
CC gene transcription in a cell encoding MEKK, where the cell is involved
CC in inflammation, regulation of cellular proliferation and
CC differentiation, regulation of development, regulation of cell death or
CC regulation of inflammation. They are also used to prepare antibodies.
CC MEKK polynucleotides can be used to produce the protein recombinantly
CC and as a source of probes and primers.
XX
SQ Sequence 626 AA;
Query Match 90.3%; Score 28; DB 20; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRYEDL 6
Db 64 VRYEDV 69
RESULT 47
AAW73533
ID AAW73533 standard; Protein: 626 AA.
XX AAW73533;
AC AAW73533;
XX 04-MAR-1999 (first entry)
DT MEKK3 protein.
DE MEKK3 protein.
XX Mitogen activated protein kinase kinase kinase; MEKK; MAPK; MEK; cancer;
KW apoptosis regulator; autoimmune disease; inflammation; allergy; therapy;
KW neuronal disorder.
XX Mammalia.
OS
XX US5854043-A.
PN 23-DEC-1998.
XX 14-OCT-1994; 94US-0323460.
XX 14-OCT-1994; 94US-0323460.
PR 14-OCT-1994; 94US-0323460.
PR 15-APR-1993; 93US-0049254.
PR 15-APR-1994; 94WO-US04178.
XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
PA Johnson GL;
PI WPI: 1999-094912/08.
DR N-PSDB; AAV45616.
XX Mitogen activated protein kinase kinase kinases and their fragments -
PT used for regulating signalling from growth factor receptors, e.g. to
PT modulate apoptosis for treatment of cancer, autoimmune disease and
PT inflammation
XX Claim 5; Column 21-26; 96pp; English.
XX This sequence is the MEKK3 protein of the invention. MEKK proteins
CC are mitogen-activated protein kinase (MAPK) kinase (MEK) kinase proteins.
CC MEKKs phosphorylate and activate MEK proteins and other signal
CC transduction molecules, so can regulate signalling initiated from a

```

CC growth factor receptor in a way different from that involving Raf
 CC protein. Particularly MEKs, or their fragments, are involved in
 CC regulation of apoptosis so they, or agents that increase their activity,
 CC are used to treat cancers, autoimmune diseases, inflammation, allergies,
 CC neuronal disorders (e.g. Alzheimer's or Parkinson's diseases) and in
 CC wound healing. MEKs are also useful for identifying agents that
 CC regulate signal transduction from cell surface receptors (e.g. from their
 CC effect on ability of MEK to phosphorylate a substrate such as MEK or Jun
 CC extracellular signal-regulated kinase).

XX Sequence 626 AA;

Query Match 90.3%; Score 28; DB 20; Length 626;
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||:
 DB 64 VRYEDV 69

RESULT 48

ID AAB01218 standard; Protein; 626 AA.

XX AAB01218;

DT 19-DEC-2000 (first entry)

DE Murine MEK3.

KW Mouse; MEK3; mitogen ERK kinase kinase; signal transduction pathway;

KW apoptosis; cancer; autoimmune disease; inflammatory response;

KW allergic response; neuronal disorder; Parkinson's disease;

XX Alzheimer's disease.

OS Mus sp.

PN US6074861-A.

PD 13-JUN-2000.

XX 05-JUN-1995; 95US-0461145.

PR 15-APR-1993; 93US-0049254.

PR 12-MAY-1995; 93US-0440421.

PR 15-APR-1994; 94WO-0504178.

PR 14-OCT-1994; 94US-0323460.

PR 14-OCT-1994; 94WO-0511690.

PR 21-FEB-1995; 95US-0354516.

XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX Johnson GL;

DR MPI; 2000-411281/35.

DR N-PSDB; AAAA9224.

XX Novel mitogen extracellular signal-regulated kinase kinase (MEK3)

PT Protein useful for treating cancer, inflammation, autoimmune diseases,

CC neurological disorders and hormone related disease in animals

PS Claim 1; Column 69-74; 92pp; English.

XX The present sequence is the murine mitogen ERK kinase kinase (MEK3)

CC protein sequence. MEK3 is involved in a signal transduction pathway

CC which can ultimately lead to apoptosis. The proteins regulated by MEK3

CC include the MEK and MAPK proteins and c-Myc. Its coding sequence was

CC identified by amplifying genomic DNA with a similar sequence to

CC the MEK3 gene. The gene and protein can be used to treat diseases such

CC as cancer, autoimmune disease, inflammatory responses, allergic responses

CC and neuronal disorders including Parkinson's disease and Alzheimer's

CC disease.

XX Sequence 626 AA;

Query Match 90.3%; Score 28; DB 21; Length 626;
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||:
 DB 64 VRYEDV 69

RESULT 49
 ID AAM48938 standard; Protein; 626 AA.

XX AAM48938;

DT 19-APR-2002 (first entry)

DE Murine MEK3.

KW Mouse; MEK3; mitogen ERK kinase kinase; enzyme; cancer; neuroprotective;

KW autoimmune disease; signal transduction; allergy; inflammation;

KW neurological disorder; hormone-related disease; apoptosis; infection;

KW cytostatic; immunosuppressive; antiinflammatory; antiallergic;

KW nootropic; antiparkinsonian; contraceptive.

OS Mus musculus.

PN US6333170-B1.

PD 25-DEC-2001.

XX 05-APR-1996; 96US-0628829.

PR 15-APR-1993; 93US-0049254.

PR 14-OCT-1994; 94US-0323460.

PR 12-MAY-1995; 95US-0440421.

PR 06-JUN-1995; 95US-0472934.

XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX Johnson GL;

DR MPI; 2002-163179/21.

DR N-PSDB; AAL41581.

XX New isolated nucleic acid encoding mitogen extracellular

PT signal-regulated kinase kinase, useful for gene therapy of e.g. cancer

PT and for recombinant protein production

XX Claim 10; Column 151-154; 125pp; English.

XX The present invention provides the protein and coding sequences of a

CC number of murine mitogen extracellular signal-regulated kinase (ERK)

CC kinase kinase (MEK3) enzymes. The sequences can be used to treat a wide

CC range of diseases including cancer, autoimmune diseases, inflammation,

CC allergies, degenerative neurological diseases and hormone-related

CC diseases, and for inhibiting spermatogenesis or oocyte maturation for

CC contraception. The present sequence is the murine MEK3 protein

CC sequence.

XX Query Match 90.3%; Score 28; DB 23; Length 626;
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||:
 DB 64 VRYEDV 69

```

RESULT 50
ID ABB68414 standard; Protein; 793 AA.
XX
AC ABB68414;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32034.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR N-PSDB: ABL12517.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure: SEQ ID NO 32034; 21pp + Sequence listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
CC sequences (AB57737-AB72072).
CC (AB57737-AB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 793 AA:

Query Match          90.3%; Score 28; DB 22; Length 793;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAREDL 6
   :|||||
DB 673 LRYEDL 678

```

Search completed: February 20, 2003, 13:32:28
 Job time : 35.6286 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:11 ; Search time 9.6 Seconds
(without alignments)

60.084 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	281	2 B90167	conserved hypothet
2	31	100.0	428	2 G75453	hypothetical prote
3	31	100.0	484	2 JE0261	N-acetylglucosamin
4	31	100.0	631	2 A64578	conserved hypothet
5	31	100.0	748	2 T30634	hypothetical prote
6	31	100.0	973	2 T21463	hypothetical prote
7	31	100.0	6260	2 T30228	polyketide synthas
8	30	96.8	562	2 JC7628	glucoside 3-dehydr
9	30	96.8	579	2 H87451	oxidoreductase, GM
10	30	96.8	1322	2 B71440	hypothetical prote
11	29	93.5	234	2 H87385	hypothetical prote
12	29	93.5	355	2 F96020	probable iron ABC
13	29	93.5	393	2 T08203	polygalacturonase
14	29	93.5	405	2 D84871	probable polygalac
15	29	93.5	430	2 G64446	ATP-dependent 26S
16	29	93.5	1401	2 T39225	MAP kinase kinase
17	29	93.5	1416	2 D71350	probable DNA-dirac
18	28	90.3	195	1 B69441	conserved hypothet
19	28	90.3	233	2 C86849	regulatory protein
20	28	90.3	273	2 T42928	immediate-early pr
21	28	90.3	277	2 C69608	chitosanase can -
22	28	90.3	306	2 A83634	hypothetical prote
23	28	90.3	333	2 A84523	probable steroid s
24	28	90.3	337	2 C38453	anaerobic sulfite
25	28	90.3	337	2 A60825	anaerobic sulfite
26	28	90.3	346	2 T29003	hypothetical prote
27	28	90.3	388	2 G70729	hypothetical prote
28	28	90.3	436	2 AH2447	molybdopterin bios
29	28	90.3	458	2 C75180	tlld related prote

30	28	90.3	458	2 A57397	Chondroitin 6-sulf
31	28	90.3	484	2 JC7350	N-acetylglucosamin
32	28	90.3	486	2 JC7351	N-acetylglucosamin
33	28	90.3	503	2 F90169	hypothetical prote
34	28	90.3	507	2 E89804	alkyl hydroperoxid
35	28	90.3	612	2 T45641	beta-D-glucan exch
36	28	90.3	1107	2 B91271	probable periplasm
37	28	90.3	1107	2 B86112	probable periplasm
38	28	90.3	1107	2 E65226	genome polyprotein
39	28	90.3	2233	1 ZLN2P3	hypothetical prote
40	27	87.1	73	2 H83960	hypothetical prote
41	27	87.1	108	2 F72649	response regulator
42	27	87.1	133	2 H69433	myosin ICI catalyt
43	27	87.1	160	2 A25571	hypothetical prote
44	27	87.1	200	2 C95315	Multi/nudix family
45	27	87.1	203	2 C95194	mutator protein [i
46	27	87.1	225	2 H98060	Mg(2+) transport A
47	27	87.1	225	2 G70340	glycoprotein 30 -
48	27	87.1	225	2 S28679	hypothetical prote
49	27	87.1	265	2 C90026	hypothetical prote
50	27	87.1	273	2 S07014	hypothetical prote
51	27	87.1	273	2 JN0037	hypothetical 32k p
52	27	87.1	283	2 E81895	probable prolipopr
53	27	87.1	307	2 E95934	probable enzyme, C
54	27	87.1	309	2 JC5697	placental transfor
55	27	87.1	311	2 D69881	yabc protein homol
56	27	87.1	312	2 B72708	probable signal re
57	27	87.1	332	2 A70772	probable sulfate A
58	27	87.1	334	2 C84130	oxidoreductase BH3
59	27	87.1	341	2 G82014	dtDPglucose 4,6-de
60	27	87.1	346	2 S47045	dtDPglucose 4,6-de
61	27	87.1	352	2 H83975	hypothetical prote
62	27	87.1	355	2 G81242	dtDPglucose 4,6-de
63	27	87.1	356	2 S49011	yolk protein 2 - b
64	27	87.1	360	2 S42431	dtDPglucose 4,6-de
65	27	87.1	386	2 F88694	protein skn-1 (imp
66	27	87.1	396	2 F89850	3-phosphoglycerate k
67	27	87.1	428	2 A69085	3-isopropylmalate
68	27	87.1	433	2 B69435	aconitase (acn) ho
69	27	87.1	439	2 H83842	diaminopimelate de
70	27	87.1	439	2 C65072	Guanine deaminase
71	27	87.1	439	2 D91098	hypothetical prote
72	27	87.1	439	2 H85943	hypothetical prote
73	27	87.1	465	2 AE3102	glutamyl-tRNA amid
74	27	87.1	465	2 E98184	probable arylamide P
75	27	87.1	474	2 B82227	exodeoxyribonuclea
76	27	87.1	476	2 H82177	conserved hypothet
77	27	87.1	488	2 D86696	SOS response unuc
78	27	87.1	488	2 T43124	conserved hypothet
79	27	87.1	499	2 S43324	zeta-carotene desa
80	27	87.1	499	2 AG2509	methy-accepting c
81	27	87.1	504	2 H87422	petriplasmic oligop
82	27	87.1	514	2 B64055	hypothetical prote
83	27	87.1	520	2 T21462	hypothetical prote
84	27	87.1	525	2 A83013	conserved hypothet
85	27	87.1	527	2 F69378	conserved hypothet
86	27	87.1	533	2 A42143	skn-1 - Caenorhabd
87	27	87.1	554	2 T51213	hypothetical prote
88	27	87.1	557	2 B86466	hypothetical prote
89	27	87.1	561	2 E98192	probable oxidoredu
90	27	87.1	561	2 AE3094	oxidoreductase Atu
91	27	87.1	562	2 A86773	hypothetical prote
92	27	87.1	605	2 G85651	probable membrane
93	27	87.1	605	2 D90791	conserved hypothet
94	27	87.1	617	2 G70039	conserved phosphoen
95	27	87.1	621	2 B71457	glucose inhibited
96	27	87.1	628	2 A64930	hypothetical prote
97	27	87.1	630	2 H97562	hypothetical prote
98	27	87.1	630	2 AH2783	peptidyl-prolyl ci
99	27	87.1	631	2 E71933	hypothetical prote
100	27	87.1	653	2 T03102	semaphorin homolog
101	27	87.1	666	2 AH1167	ABC transporter, A
102	27	87.1	685	2 F96032	hypothetical prote

103	27	87.1	694	2	D86615	176	26	83.9	686	2	D71292	probable DNA recom
104	27	87.1	694	2	G72009	177	26	83.9	686	2	B96526	unknown protein (1
105	27	87.1	716	2	T37710	178	26	83.9	693	1	B64139	DNA helicase recog
106	27	87.1	747	2	T42599	179	26	83.9	721	2	B87611	tonB-dependent rec
107	27	87.1	753	1	WZBER8	180	26	83.9	727	1	S17854	NADH2 dehydrogenas
108	27	87.1	759	1	EB7443	181	26	83.9	728	2	S45403	hypothetical prote
109	27	87.1	769	1	WZBE54	182	26	83.9	760	2	G71417	hypothetical prote
110	27	87.1	769	1	B87681	183	26	83.9	762	2	JC7179	N,N-dimethylformam
111	27	87.1	773	2	T46283	184	26	83.9	799	2	F83456	xanthine dehydroge
112	27	87.1	785	2	B72608	185	26	83.9	830	2	F83288	conserved hypotet
113	27	87.1	809	2	B87455	186	26	83.9	849	2	E86306	Similar to tutei1
114	27	87.1	891	2	T40417	187	26	83.9	886	1	UC5085	replication licens
115	27	87.1	926	2	T04679	188	26	83.9	892	1	S44228	replication licens
116	27	87.1	976	2	A87319	189	26	83.9	904	1	T10067	replication licens
117	27	87.1	1060	2	F68710	190	26	83.9	975	2	AC2517	hypothetical prote
118	27	87.1	1079	2	T30996	191	26	83.9	975	2	A86258	protein F5011.4 [i
119	27	87.1	1094	2	S49313	192	26	83.9	1057	1	OYMSAR	atrial natriuretic
120	27	87.1	1096	1	S61917	193	26	83.9	1057	1	OCMSAR	atrial natriuretic
121	27	87.1	1137	2	T19414	194	26	83.9	1057	2	T57963	natriuretic peptid
122	27	87.1	1174	2	T43081	195	26	83.9	1057	2	T55319	guanylyl cyclase A
123	27	87.1	1352	2	G71051	196	26	83.9	1166	2	T15628	hypothetical prote
124	27	87.1	1359	2	T10235	197	26	83.9	1226	2	AB3327	hypothetical membr
125	27	87.1	1393	2	E72122	198	26	83.9	1396	2	G71529	DNA-directed RNA p
126	27	87.1	1393	2	B86501	199	26	83.9	1396	2	F81686	DNA-directed RNA p
127	27	87.1	1397	2	B81548	200	26	83.9	1397	2	T46354	hypothetical prote
128	27	87.1	1704	2	T43141	201	26	83.9	1450	2	A84780	probable ABC trans
129	27	87.1	3433	1	S28381	202	26	83.9	2548	2	E59435	myosin IXA [limpo
130	27	87.1	10223	2	T30225	203	26	83.9	2626	2	T31099	myosin-Rhocap prot
131	26	83.9	105	2	F90503	204	26	83.9	3144	2	S64791	VPS13 protein - ye
132	26	83.9	121	2	S35234	205	26	80.6	72	2	T49026	ubiquinol-cytochro
133	26	83.9	136	2	C81436	206	26	80.6	88	2	T07369	ubiquinol-cytochro
134	26	83.9	166	2	T49694	207	25	80.6	97	2	T07568	ribosomal protein
135	26	83.9	166	2	T64740	208	25	80.6	102	2	S57628	thymidylate syntha
136	26	83.9	201	2	F97063	209	25	80.6	114	2	A71126	hypothetical prote
137	26	83.9	231	2	T22350	210	25	80.6	117	1	W0BP37	flav protein [limpo
138	26	83.9	246	2	C70316	211	25	80.6	122	2	T05357	gene 0.3 protein -
139	26	83.9	253	2	D95121	212	25	80.6	133	2	B82977	conserved hypotet
140	26	83.9	253	2	H97990	213	25	80.6	137	2	T41575	ubiquinol-cytochro
141	26	83.9	261	2	S10321	214	25	80.6	145	2	F83267	conserved hypotet
142	26	83.9	261	2	S75926	215	25	80.6	147	2	F72310	hypothetical prote
143	26	83.9	284	2	T22671	216	25	80.6	162	2	D69899	conserved hypotet
144	26	83.9	288	1	B48583	217	25	80.6	163	2	AC1939	peptide methionine
145	26	83.9	302	2	D83479	218	25	80.6	163	2	H96014	conserved hypotet
146	26	83.9	306	2	T29990	219	25	80.6	169	2	A64776	probable membrane
147	26	83.9	339	2	T28178	220	25	80.6	169	2	G90692	probable membrane
148	26	83.9	344	2	A95053	221	25	80.6	185	2	C85543	conserved hypotet
149	26	83.9	356	2	F97923	222	25	80.6	200	2	H97513	nucol (A7245398) [
150	26	83.9	356	2	AC2907	223	25	80.6	200	2	AC2732	NADH ubiquinone ox
151	26	83.9	365	2	B69220	224	25	80.6	203	2	AC3611	formylmethionine d
152	26	83.9	378	2	G97682	225	25	80.6	208	2	AC3290	transcription regu
153	26	83.9	384	2	G82976	226	25	80.6	210	2	AD1735	hypothetical prote
154	26	83.9	408	2	T40698	227	25	80.6	238	2	A69463	2-hydroxy-6-oxo-6-
155	26	83.9	413	2	T08297	228	25	80.6	241	2	B97164	sigma factor of St
156	26	83.9	430	2	S73475	229	25	80.6	248	2	S34753	epithelial cell ma
157	26	83.9	448	2	S45112	230	25	80.6	258	2	H72269	ABC transporter, A
158	26	83.9	458	1	E71033	231	25	80.6	264	2	G75045	probable zinc-bind
159	26	83.9	465	2	A31810	232	25	80.6	286	2	A72583	hypothetical prote
160	26	83.9	488	2	B96521	233	25	80.6	286	2	T29719	hypothetical prote
161	26	83.9	498	2	S12061	234	25	80.6	300	1	S16815	SNP1 protein - yea
162	26	83.9	507	2	C82901	235	25	80.6	305	2	G95122	ABC transporter, A
163	26	83.9	526	1	P5XR15	236	25	80.6	310	2	A33489	probable transcrip
164	26	83.9	526	1	P5XR15	237	25	80.6	316	1	T27194	hypothetical prote
165	26	83.9	526	1	A45185	238	25	80.6	318	1	YKSA73	thymidylate syntha
166	26	83.9	526	1	S18768	239	25	80.6	318	2	C89920	thymidylate syntha
167	26	83.9	527	1	S18762	240	25	80.6	321	2	H71729	hypothetical prote
168	26	83.9	527	1	G97133	241	25	80.6	326	2	H82987	magnesium/cobalt t
169	26	83.9	596	2	S36475	242	25	80.6	331	2	D66515	hypothetical prote
170	26	83.9	602	2	H82635	243	25	80.6	331	2	T25184	
171	26	83.9	615	2	F70783	244	25	80.6				
172	26	83.9	633	2	E86509	245	25	80.6				
173	26	83.9	662	2	B72114	246	25	80.6				
174	26	83.9	662	2	T23271	247	25	80.6				
175	26	83.9	671	2	E82355	248	25	80.6				

249	25	80.6	334	2	S74737	conserved hypotnet	322	25	80.6	664	2	T12988	hypothetical prote
250	25	80.6	346	2	B69350	conserved hypotnet	323	25	80.6	716	2	T37830	probable glutamate
251	25	80.6	346	2	E70904	hypothetical prote	324	25	80.6	727	2	A33552	NADH2 dehydrogenas
252	25	80.6	347	2	AC3840	alcohol dehydrogen	325	25	80.6	736	2	T05137	protein kinase
253	25	80.6	350	2	C64369	endoglucanase homo	326	25	80.6	778	2	JC7797	scallop unconvetit
254	25	80.6	350	2	D70805	hypothetical prote	327	25	80.6	787	2	C75058	probable beta-gala
255	25	80.6	351	2	F90409	GTP binding conser	328	25	80.6	801	2	A11960	hypothetical prote
256	25	80.6	357	2	A97618	alcohol dehydrogen	329	25	80.6	866	2	C97662	hypothetical prote
257	25	80.6	363	2	S75088	probable GTP-bind	330	25	80.6	866	2	AF2886	conserved hypotnet
258	25	80.6	365	2	S74847	hypothetical prote	331	25	80.6	873	2	T25442	hypothetical prote
259	25	80.6	366	2	S66016	probable GTP-bind	332	25	80.6	883	1	RMBP17	DNA-directed RNA p
260	25	80.6	370	2	C69309	conserved hypotnet	333	25	80.6	884	1	RMBP13	DNA-directed RNA p
261	25	80.6	372	2	F70467	hypothetical prote	334	25	80.6	897	2	T43628	phosphatidylinosit
262	25	80.6	374	2	AE2227	transposase alr337	335	25	80.6	899	2	S12319	pre-mRNA splicing
263	25	80.6	375	1	A23689	limulus clotting e	336	25	80.6	928	2	S50578	hypothetical prote
264	25	80.6	377	2	S21302	succinate dehydrog	337	25	80.6	929	2	T35683	ftsK homolog - Str
265	25	80.6	385	2	S46532	polygalacturonase	338	25	80.6	971	2	H71719	hypothetical prote
266	25	80.6	394	2	F97178	spore coat polysac	339	25	80.6	976	2	H97269	zn-dependent pepti
267	25	80.6	400	2	T25605	hypothetical prote	340	25	80.6	979	2	J00894	Pil5 protein - Myc
268	25	80.6	402	2	T45518	hypothetical prote	341	25	80.6	979	2	B80601	hypothetical prote
269	25	80.6	402	2	AC2304	transposase alr388	342	25	80.6	988	2	S77211	isooleucine-tRNA Ii
270	25	80.6	402	2	AC2552	transposase alr1801	343	25	80.6	1013	2	B75583	hypothetical prote
271	25	80.6	402	2	AD2506	transposase alr722	344	25	80.6	1052	2	T00067	hypothetical prote
272	25	80.6	402	2	AC2506	transposase alr723	345	25	80.6	1095	2	PC1114	SKDC25 protein -
273	25	80.6	402	2	AI2318	transposase alr410	346	25	80.6	1133	2	T01920	probable RNA-direc
274	25	80.6	402	2	H86649	thiamosyltransfera	347	25	80.6	1249	2	S54376	tripeptidyl-peptid
275	25	80.6	403	2	AE2518	transposase alr732	348	25	80.6	1249	2	S68431	tripeptidyl-peptid
276	25	80.6	403	2	AI2363	transposase alr446	349	25	80.6	1262	2	I48835	tripeptidyl-peptid
277	25	80.6	403	2	AB1951	transposase alr115	350	25	80.6	1286	2	T18734	hypothetical prote
278	25	80.6	403	2	B95175	TPR domain protein	351	25	80.6	1302	2	T23236	hypothetical prote
279	25	80.6	409	1	EFEET	translation elonga	352	25	80.6	1450	2	T45888	ABC transporter-11
280	25	80.6	413	2	AG2456	transposase alr520	353	25	80.6	1465	2	S31262	TYB protein - years
281	25	80.6	414	2	C98041	conserved hypotnet	354	25	80.6	1467	2	PC1253	TYB protein - years
282	25	80.6	425	2	H90415	hypothetical prote	355	25	80.6	1490	2	UC5145	DNA (cytosine-5')-
283	25	80.6	436	2	D84782	probable proline t	356	25	80.6	1537	2	UC4172	DNA (cytosine-5')-
284	25	80.6	439	2	T47713	proline transporte	357	25	80.6	1676	2	E71410	probable centromer
285	25	80.6	439	2	T50688	proline transport	358	25	80.6	1680	2	T01367	hypothetical prote
286	25	80.6	441	2	T50689	proline transport	359	25	80.6	1690	2	S41467	DNA-directed RNA p
287	25	80.6	442	2	T50692	proline transport	360	25	80.6	1748	2	S63127	probable membrane
288	25	80.6	442	2	T50687	proline transport	361	25	80.6	1802	2	S52611	TYB protein - years
289	25	80.6	442	2	C82823	type I restriction	362	25	80.6	1803	2	S68894	TYB protein - years
290	25	80.6	444	2	T20600	hypothetical prote	363	25	80.6	1819	2	T12008	hypothetical prote
291	25	80.6	450	2	G83536	hypothetical prote	364	25	80.6	1819	2	T17428	FK506 polypeptide s
292	25	80.6	452	2	A70389	L-seryl-tRNAsec se	365	25	80.6	7576	2	E72588	hypothetical prote
293	25	80.6	456	2	T19817	hypothetical prote	366	25	80.6	7576	2	J50654	hypothetical 8.2K
294	25	80.6	471	2	AC3622	mannose-1-phosphat	367	25	80.6	83	2	E64482	hypothetical prote
295	25	80.6	473	2	T50690	proline transport	368	25	80.6	86	2	T12077	late embryogenesis
296	25	80.6	481	2	T01850	UTP-glucose glucos	369	25	80.6	89	2	D82298	ribosomal protein
297	25	80.6	488	2	F71825	replicative DNA he	370	25	80.6	91	2	H82331	conserved hypotnet
298	25	80.6	488	2	T09734	1-aminocyclopropan	371	25	80.6	100	2	S69860	hypothetical prote
299	25	80.6	488	2	B64690	replicative DNA he	372	25	80.6	102	2	JC2343	heat shock protein
300	25	80.6	491	2	E69368	glutamine syntheta	373	25	80.6	102	2	JC2344	heat shock protein
301	25	80.6	498	2	B75373	GGDEF family prote	374	25	80.6	102	2	A72739	probable DNA-direc
302	25	80.6	500	2	T16630	hypothetical prote	375	25	80.6	110	1	VHBPDL	major capsid prote
303	25	80.6	520	1	S50990	galactokinase (EC	376	25	80.6	110	2	B80833	major capsid prote
304	25	80.6	532	2	S54571	probable membrane	377	25	80.6	110	2	G30900	probable head deco
305	25	80.6	535	2	A84464	hypothetical prote	378	25	80.6	114	2	G89566	protein T08A9.6 [l
306	25	80.6	536	2	F90299	acylaminoacyl-pept	379	25	80.6	115	2	D85690	hypothetical prote
307	25	80.6	548	2	S27958	transcription fact	380	25	80.6	115	2	B70950	hypothetical prote
308	25	80.6	550	2	H71952	ATP-dependent zinc	381	25	80.6	117	2	G90464	hypothetical prote
309	25	80.6	550	2	F64555	cell division prot	382	25	80.6	121	2	D48234	hypothetical prote
310	25	80.6	558	2	D88163	protein F10C1.2a [383	25	80.6	125	2	T38537	probable single-st
311	25	80.6	573	2	S45903	hypothetical prote	384	25	80.6	125	2	S46670	bulious pemphigoid
312	25	80.6	575	2	AD2361	flavoprotein [limpo	385	25	80.6	126	2	C97616	hypothetical prote
313	25	80.6	585	2	S48929	hypothetical prote	386	25	80.6	126	2	A12838	hypothetical prote
314	25	80.6	588	2	H89885	succinate dehydrog	387	25	80.6	127	2	AB2252	aspartate 1-decarb
315	25	80.6	589	2	S46326	hypothetical prote	388	25	80.6	127	2	B89862	conserved hypotnet
316	25	80.6	595	2	H87350	dehydratase, ilvd/	389	25	80.6	129	2	A30798	lysine carboxypept
317	25	80.6	603	2	T04733	auxin-regulated pr	390	25	80.6	130	2	H72722	hypothetical prote
318	25	80.6	605	2	S36592	El protein - human	391	25	80.6	131	2	AB1586	transcription regu
319	25	80.6	609	2	S36481	ORC3 protein - yea	392	25	80.6	132	2	B84295	hypothetical prote
320	25	80.6	616	2	S64746	conserved hypotnet	393	25	80.6	139	2	T33713	conserved hypotnet
321	25	80.6	620	2	S64304	hypothetical prote	394	25	80.6	141	2	T28189	hypothetical prote

395	24	77.4	142	2	B86552	glycerol-3-phospha
396	24	77.4	147	2	T35563	ribosomol protein
397	24	77.4	150	2	T36372	probable asnc-fam1
398	24	77.4	151	2	C90048	conserved hypotnet
399	24	77.4	153	2	T16012	hypothetical prote
400	24	77.4	156	2	A41509	myosin essential 1
401	24	77.4	156	2	UT0902	chaperonin 60 beta
402	24	77.4	157	2	A25183	myosin essential 1
403	24	77.4	158	2	H87643	conserved hypotnet
404	24	77.4	164	2	A81335	probable signal-tr
405	24	77.4	165	2	G69304	transcription regu
406	24	77.4	165	2	G63415	hypothetical prote
407	24	77.4	169	2	T21917	3'-aminoacylcoside
408	24	77.4	170	2	I40612	hypothetical prote
409	24	77.4	172	2	JEO381	NADH2 dehydrogenas
410	24	77.4	174	2	B75176	hypothetical prote
411	24	77.4	177	2	PS0374	chaperonin 60 beta
412	24	77.4	177	2	S71014	hypothetical prote
413	24	77.4	177	2	C70679	probable atpd prot
414	24	77.4	183	2	A84527	hypothetical prote
415	24	77.4	184	2	T28333	ORF MSV172 hypote
416	24	77.4	185	2	E70445	C-terminal fumarat
417	24	77.4	188	2	E72860	ORF-84 protein - A
418	24	77.4	190	1	K1SSMA	calcium-binding pr
419	24	77.4	190	2	A83102	conserved hypotnet
420	24	77.4	190	2	S48062	NADH2 dehydrogenas
421	24	77.4	190	2	T03231	NADH2 dehydrogenas
422	24	77.4	192	1	K1SSAB	calcium-binding pr
423	24	77.4	192	2	S62130	NADH2 dehydrogenas
424	24	77.4	192	2	T11930	NADH2 dehydrogenas
425	24	77.4	193	2	D64769	YajB protein - Esc
426	24	77.4	193	2	G90685	probable glycoprot
427	24	77.4	193	2	C85356	probable glycoprot
428	24	77.4	193	2	H90474	hypothetical prote
429	24	77.4	194	2	T48071	hypothetical prote
430	24	77.4	196	2	A41163	Ca2+/calmodulin-de
431	24	77.4	197	2	E87491	NADH dehydrogenase
432	24	77.4	199	2	A32183	tropomyosin TPM1 -
433	24	77.4	199	2	S75344	hypothetical prote
434	24	77.4	200	2	H98209	hypothetical prote
435	24	77.4	200	2	A13076	conserved hypotnet
436	24	77.4	200	2	F90098	putative ribosome
437	24	77.4	201	2	F70468	conserved hypotnet
438	24	77.4	203	2	G69898	conserved hypotnet
439	24	77.4	204	2	S78165	NADH2 dehydrogenas
440	24	77.4	208	2	F90223	conserved hypotnet
441	24	77.4	210	2	G97233	hypothetical prote
442	24	77.4	212	2	AC1055	peptide methionine
443	24	77.4	212	2	S25996	hypothetical prote
444	24	77.4	213	2	B84333	isopentenyl pyroph
445	24	77.4	213	2	C70943	hypothetical prote
446	24	77.4	215	1	ADECFP	L-fuculose-1-phosp
447	24	77.4	215	2	D91086	L-fuculose-1-phosp
448	24	77.4	215	2	F85931	N-acylhomoserine 1
449	24	77.4	216	2	A10120	probable heme expo
450	24	77.4	218	2	AF0333	hypothetical prote
451	24	77.4	227	2	T49725	hypothetical prote
452	24	77.4	228	2	T16678	hypothetical prote
453	24	77.4	232	2	AD0071	probable DedA-fam1
454	24	77.4	232	2	E70372	hypothetical prote
455	24	77.4	233	2	A81366	protein gp33 [Bact
456	24	77.4	233	2	A13590	CD30 ligand - huma
457	24	77.4	234	2	A40710	gp53 protein - Myc
458	24	77.4	235	2	B72806	hypothetical prote
459	24	77.4	237	2	AE2037	thiol-disulfide in
460	24	77.4	238	2	C70479	hypothetical prote
461	24	77.4	241	2	B83907	hypothetical prote
462	24	77.4	242	2	A71972	conserved hypotnet
463	24	77.4	242	2	C64537	minor tail protein
464	24	77.4	243	2	T13105	hypothetical prote
465	24	77.4	244	2	S67186	merr-family transc
466	24	77.4	244	2	A11414	photosystem II oxy
467	24	77.4	245	2	S00413	
468	24	77.4	468	2	S66691	neutrophil granule
469	24	77.4	469	2	S03888	photosystem II oxy
470	24	77.4	470	2	T02122	hypothetical prote
471	24	77.4	471	2	C83865	hypothetical prote
472	24	77.4	472	2	G97340	acyl-ACP thioester
473	24	77.4	473	2	E84072	hypothetical prote
474	24	77.4	474	2	D69363	conserved hypotnet
475	24	77.4	475	2	T03873	photosystem II oxy
476	24	77.4	476	2	H90656	probable membrane
477	24	77.4	477	2	H85507	probable membrane
478	24	77.4	478	2	T02873	conserved phototyst
479	24	77.4	479	2	F90408	conserved hypotnet
480	24	77.4	480	2	AG0437	probable membrane
481	24	77.4	481	2	D86693	conserved hypotnet
482	24	77.4	482	2	G65206	NADH pyrophosphata
483	24	77.4	483	2	G91243	hypothetical prote
484	24	77.4	484	2	E86091	hypothetical prote
485	24	77.4	485	2	AD0932	heat-labile entero
486	24	77.4	486	2	OLECA	photosystem II oxy
487	24	77.4	487	2	F210X2	photosystem II oxy
488	24	77.4	488	2	S22763	photosystem II oxy
489	24	77.4	489	2	J50771	photosystem II oxy
490	24	77.4	490	2	A75547	hypothetical prote
491	24	77.4	491	2	A72116	biotin apo-protein
492	24	77.4	492	2	A86507	biotin protein lig
493	24	77.4	493	2	S10016	photosystem II oxy
494	24	77.4	494	2	T32440	hypothetical prote
495	24	77.4	495	2	J70390	type II site-speci
496	24	77.4	496	2	YAZGF7	S-antigen precurs
497	24	77.4	497	2	S63532	NAD(P)H-quinone ox
498	24	77.4	498	2	S00005	photosystem II oxy
499	24	77.4	499	2	S17446	photosystem II oxy
500	24	77.4	500	2	T00088	ABC-type transport
501	24	77.4	501	2	F86482	protein F5J5.7 (lm
502	24	77.4	502	2	PRECT9	kanamycin kinase (
503	24	77.4	503	2	J00103	kanamycin kinase (
504	24	77.4	504	2	J01545	kanamycin kinase (
505	24	77.4	505	2	S16630	APH protein - Salm
506	24	77.4	506	2	H83525	kanamycin kinase (
507	24	77.4	507	2	E86833	conserved hypotnet
508	24	77.4	508	2	T03942	transcription regu
509	24	77.4	509	2	S75709	rRNA N-glycosidase
510	24	77.4	510	2	S24723	hypothetical prote
511	24	77.4	511	2	B82323	tryptophan synthas
512	24	77.4	512	2	S39072	dimethyladenosine
513	24	77.4	513	2	C97742	NADH2 dehydrogenas
514	24	77.4	514	2	B90256	hypothetical prote
515	24	77.4	515	2	S74914	conserved hypotnet
516	24	77.4	516	2	G71300	hypothetical prote
517	24	77.4	517	2	T03085	probable protein-m
518	24	77.4	518	2	T03122	ribonuclease homol
519	24	77.4	519	2	S51571	hypothetical prote
520	24	77.4	520	2	E75366	hypothetical prote
521	24	77.4	521	2	A53372	glutanyl-tRNA synt
522	24	77.4	522	2	E69025	lysR-type regulato
523	24	77.4	523	2	C70251	conserved hypotnet
524	24	77.4	524	2	C70231	hypothetical prote
525	24	77.4	525	2	T36255	hypothetical prote
526	24	77.4	526	2	R1ZMR7	hypothetical prote
527	24	77.4	527	2	F82287	rRNA N-glycosidase
528	24	77.4	528	2	S11860	transcription regu
529	24	77.4	529	2	G83093	rRNA N-glycosidase
530	24	77.4	530	2	S11859	UDP-3-O-acyl-N-ace
531	24	77.4	531	2	A83239	rRNA N-glycosidase
532	24	77.4	532	2	A83239	conserved hypotnet
533	24	77.4	533	2	YAKON7	S-antigen precurs
534	24	77.4	534	2	F71976	probable lacyl-car
535	24	77.4	535	2	B64531	probable lacyl-car
536	24	77.4	536	2	F90540	hypothetical prote
537	24	77.4	537	2	T36133	cysteine synthase
538	24	77.4	538	2	G70373	Mg(2+) and Co(2+)
539	24	77.4	539	2	D90482	dehydrogenase, pro
540	24	77.4	540	2	H87443	NADP-dependent qui
	24	77.4	540	2	A42507	F5L protein - vac

687	24	77.4	529	2	T33738	hypothetical prote	760	24	77.4	727	2	G84487	probable DNA repli
688	24	77.4	530	2	T42520	conserved hypochet	761	24	77.4	728	1	HHC890	heat shock protein
689	24	77.4	531	2	B55788	glutaminosyltransf	762	24	77.4	730	2	D72210	conserved hypochet
690	24	77.4	536	2	S66716	glutamate-tRNA lig	763	24	77.4	732	1	S05238	peptidyl-dipeptida
691	24	77.4	542	2	H64339	hypothetical prote	764	24	77.4	732	1	HHH886	heat shock protein
692	24	77.4	543	2	T35352	probable proteinase	765	24	77.4	732	1	HHC895	hypothetical prote
693	24	77.4	543	2	S64850	probable membrane	766	24	77.4	733	1	HHC886	hypothetical prote
694	24	77.4	544	1	S41389	penton protein (II	767	24	77.4	737	2	T46243	hypothetical prote
695	24	77.4	544	2	H87473	medium-chain-fatty	768	24	77.4	764	2	T45793	hypothetical prote
696	24	77.4	545	2	S27995	probable acid-CoA	769	24	77.4	766	2	T20003	hypothetical prote
697	24	77.4	549	2	JC2564	heat shock protein	770	24	77.4	775	1	A36631	replication licens
698	24	77.4	550	1	A34576	crystal protein pr	771	24	77.4	776	2	T48317	cyclin F mouse
699	24	77.4	559	2	H89931	DNA repair protein	772	24	77.4	779	2	G87573	xanthine dehydrog
700	24	77.4	559	2	H82532	conserved hypochet	773	24	77.4	782	2	G96698	hypothetical prote
701	24	77.4	561	2	G96752	unknown protein F2	774	24	77.4	783	2	T37457	tricorn proteinase
702	24	77.4	565	2	H69342	GTP-binding protei	775	24	77.4	785	2	I50180	cadherin-7 - chick
703	24	77.4	569	2	A11347	hypothetical prote	776	24	77.4	786	2	A55501	cyclin F - human
704	24	77.4	572	2	T41371	hypothetical prote	777	24	77.4	788	2	T25667	hypothetical prote
705	24	77.4	581	2	T33737	hypothetical prote	778	24	77.4	789	1	A45617	hypothetical prote
706	24	77.4	588	2	FW0007	hypothetical prote	779	24	77.4	790	2	I51638	6-phosphofructokin
707	24	77.4	589	2	H90112	Chaperonin 62.5K b	780	24	77.4	790	2	E85726	F-cadherin - Afri
708	24	77.4	589	2	S06954	Chaperonin 60 beta	781	24	77.4	790	2	B64903	hypothetical prote
709	24	77.4	590	2	D84523	intermediate filam	782	24	77.4	790	2	D90891	ydb protein precu
710	24	77.4	593	2	A48459	S antigen, heat st	783	24	77.4	792	2	B82756	hypothetical prote
711	24	77.4	595	2	T06412	probable chaperoni	784	24	77.4	805	2	B75515	organic solvent to
712	24	77.4	597	2	S74632	translation elonga	785	24	77.4	807	2	A64575	sensory box/GGDP
713	24	77.4	600	2	UT0901	Chaperonin 60 beta	786	24	77.4	807	2	T02738	VibA homolog - He
714	24	77.4	600	2	B96597	Rubisco subunit b1	787	24	77.4	812	2	A53016	probable AAA-type
715	24	77.4	601	2	T02633	hypothetical prote	788	24	77.4	830	2	C86191	myosin heavy chain
716	24	77.4	604	2	H89914	hypothetical prote	789	24	77.4	838	2	F97768	hypothetical prote
717	24	77.4	607	2	H81676	lipid-A-disacchari	790	24	77.4	843	2	H86209	hypothetical prote
718	24	77.4	610	2	S41315	hypothetical prote	791	24	77.4	862	1	A49346	protein F22G5.10 l
719	24	77.4	613	2	T43933	DNA-directed DNA p	792	24	77.4	882	2	AC2082	aldehyde dehydrog
720	24	77.4	613	2	F87390	TonB-dependent DNA	793	24	77.4	893	2	T03864	ferric aerobactin
721	24	77.4	618	2	A75015	phosphoenolpyruvat	794	24	77.4	901	2	D70116	hypothetical prote
722	24	77.4	619	2	I64087	translational elonga	795	24	77.4	908	2	C70168	transcription elon
723	24	77.4	620	2	H85431	ATPase-like protei	796	24	77.4	913	2	F82958	DNA polymerase I p
724	24	77.4	627	2	T21605	hypothetical prote	797	24	77.4	935	2	E84491	hypothetical prote
725	24	77.4	632	2	H70339	NADH2 dehydrogenas	798	24	77.4	936	2	T23393	hypothetical prote
726	24	77.4	637	2	H86693	DNA primase (EC 2.	799	24	77.4	946	1	IYHR2	inter-alpha-trypsi
727	24	77.4	638	2	D69957	conserved hypochet	800	24	77.4	960	2	A48083	chromosome segrega
728	24	77.4	638	2	AD1166	probable peptidogl	801	24	77.4	982	2	T18576	serine-threonine k
729	24	77.4	640	2	F90587	lipoprotein (Impor	802	24	77.4	1057	4	B47521	RNA-dependent RNA
730	24	77.4	642	2	JC2485	DNA primase (EC 2.	803	24	77.4	1077	2	S45395	hypothetical prote
731	24	77.4	644	2	T21137	hypothetical prote	804	24	77.4	1082	2	S37837	suppressor protein
732	24	77.4	648	2	G72279	ATP-dependent DNA	805	24	77.4	1108	2	T41188	probable ubiquitin
733	24	77.4	650	2	B87466	TonB-dependent rec	806	24	77.4	1108	2	AF1047	probable membrane
734	24	77.4	664	2	A72215	exonuclease ABC c	807	24	77.4	1112	2	D75036	cell division cont
735	24	77.4	664	2	I39467	bullous pemphigoid	808	24	77.4	1113	2	S48495	probable membrane
736	24	77.4	664	2	T17317	hypothetical prote	809	24	77.4	1184	2	D86387	probable protein p
737	24	77.4	668	1	F69794	DNA ligase (NAD) (810	24	77.4	1195	1	S26722	DNA-directed RNA p
738	24	77.4	678	2	H88187	protein C18H9.8 y1	811	24	77.4	1195	2	E96615	hypothetical prote
739	24	77.4	679	2	T40219	hypothetical prote	812	24	77.4	1206	2	S24407	formin isoform IV
740	24	77.4	680	2	H81987	probable DNA helic	813	24	77.4	1213	2	A41724	limb deformity (ld
741	24	77.4	680	2	E81043	ATP-dependent DNA	814	24	77.4	1221	2	T30529	ubiquitin carboxyl
742	24	77.4	686	1	D70172	DNA recombinase (r	815	24	77.4	1237	2	AC1583	internalin protein
743	24	77.4	688	2	S61249	probable virion pr	816	24	77.4	1265	2	T06916	DNA-directed RNA p
744	24	77.4	688	2	T48176	receptor like prot	817	24	77.4	1289	2	I84505	calcium-dependent
745	24	77.4	692	2	H82041	ATP-dependent DNA	818	24	77.4	1306	1	A31759	peritidy1-dipeptid
746	24	77.4	693	1	JH0265	DNA recombinase (E	819	24	77.4	1316	1	H1BPD7	internal virion pr
747	24	77.4	693	2	G91194	DNA helicase RecG	820	24	77.4	1331	1	XORPDH	xanthine dehydrog
748	24	77.4	693	2	AE0005	ATP-dependent DNA	821	24	77.4	1335	1	XOMSDH	xanthine dehydrog
749	24	77.4	693	2	AE0969	ATP-dependent DNA	822	24	77.4	1364	2	T10236	xanthine dehydrog
750	24	77.4	703	2	A84563	probable AAA-type	823	24	77.4	1390	2	A45455	nucleoporin 155 -
751	24	77.4	704	2	H86041	hypothetical prote	824	24	77.4	1413	2	G84790	probable ABC trans
752	24	77.4	712	2	B87683	TonB-dependent rec	825	24	77.4	1435	2	D96693	protein putative A
753	24	77.4	717	1	HHFR83	heat shock protein	826	24	77.4	1437	2	C75198	activator 1, repli
754	24	77.4	718	2	C82817	ATP-dependent DNA	827	24	77.4	1451	2	A36468	SPT protein - yea
755	24	77.4	719	2	E88504	protein B0361.8 [1	828	24	77.4	1468	2	S11515	formin - mouse
756	24	77.4	721	2	T40945	hypothetical prote	829	24	77.4	1469	2	H96622	probable ABC trans
757	24	77.4	724	1	HHH084	heat shock protein	830	24	77.4	1588	2	T38660	probable transcrip
758	24	77.4	724	1	HHMS84	heat shock protein	831	24	77.4	1688	2	D75489	hypothetical prote
759	24	77.4	725	2	JC1468	heat shock protein	832	24	77.4	1679	2	S48385	hypothetical prote

833	24	77.4	1336	2	A47747	light junction pro
834	24	77.4	1192	2	T13939	myosin V - fruit f
835	24	77.4	1828	2	B59254	myosin heavy chain
836	24	77.4	1830	1	S19188	myosin-V - chicken
837	24	77.4	1853	1	A46761	myosin heavy chain
838	24	77.4	1855	2	A59254	transcription init
839	24	77.4	1865	1	I48155	gag-pol-like fusio
840	24	77.4	1870	2	A47521	hypothetical prote
841	24	77.4	1882	1	T00069	transcription init
842	24	77.4	1893	1	A40262	genome polyprotein
843	24	77.4	2206	1	GNNY21	polymerase - Berne
844	24	77.4	2291	1	S11238	hypothetical prote
845	24	77.4	2531	2	T16743	hypothetical prote
846	24	77.4	2549	2	A40937	bulbosus pemphigoid
847	24	77.4	2748	2	S51976	nuclear migration
848	24	77.4	2810	2	T22298	hypothetical prote
849	24	77.4	2871	2	A55567	fibrillin I - bovi
850	24	77.4	3002	2	A47221	fibrillin I precur
851	24	77.4	3068	2	S69625	hypothetical prote
852	24	77.4	3432	1	GNNVJ5	genome polyprotein
853	24	77.4	3432	1	GNNVJ5	genome polyprotein
854	24	77.4	3587	1	GNNVJ5	surfactin syntheta
855	24	77.4	4630	2	E68679	polyketide synthet
856	24	77.4	7962	2	I38346	elastic titin - hu
857	23	74.2	31	2	S49191	hypothetical prote
858	23	74.2	45	2	S75733	protein YMR158c-b
859	23	74.2	48	2	F66043	unknown protein en
860	23	74.2	55	2	D33052	FlxI protein - Rhi
861	23	74.2	55	2	B95344	FlxI nitroge fix
862	23	74.2	58	2	S58628	hypothetical prote
863	23	74.2	58	2	T12949	hypothetical prote
864	23	74.2	61	2	H91196	hypothetical prote
865	23	74.2	61	2	G69257	hypothetical prote
866	23	74.2	63	1	Q0RC1M	yaIA protein - Esc
867	23	74.2	63	2	AD0550	conserved hypotbet
868	23	74.2	63	2	G85534	hypothetical prote
869	23	74.2	63	2	G90683	hypothetical prote
870	23	74.2	64	2	T36824	hypothetical prote
871	23	74.2	71	2	J00277	hypothetical bk pr
872	23	74.2	72	2	AH3281	hypothetical cytos
873	23	74.2	85	2	S31018	gene 73 protein -
874	23	74.2	87	1	R3NT15	ribosomal protein
875	23	74.2	87	2	C69990	hypothetical prote
876	23	74.2	88	2	E71358	conserved hypotbet
877	23	74.2	92	2	A87600	hypothetical prote
878	23	74.2	93	2	H90050	hypothetical prote
879	23	74.2	95	2	AF1069	hypothetical prote
880	23	74.2	96	2	BA1979	neuropeptide 1 pre
881	23	74.2	96	2	S49377	translational releas
882	23	74.2	96	2	B66701	protein Fl2A21.1 l
883	23	74.2	102	2	S57536	hypothetical prote
884	23	74.2	106	2	T46357	hypothetical prote
885	23	74.2	108	2	A13250	viralG regulated p
886	23	74.2	109	2	S12338	hypothetical prote
887	23	74.2	109	2	GC4402	hypothetical prote
888	23	74.2	110	2	A11845	hypothetical prote
889	23	74.2	113	2	DB3489	hypothetical prote
890	23	74.2	114	2	B97036	probable metal-bin
891	23	74.2	120	2	T44411	ribosomal protein
892	23	74.2	122	1	R5SP14	ribosomal protein
893	23	74.2	122	1	S24287	ig heavy chain V r
894	23	74.2	123	1	R5NT14	ribosomal protein
895	23	74.2	123	1	R5R214	ribosomal protein
896	23	74.2	123	1	R3ZM14	ribosomal protein
897	23	74.2	126	2	S72785	hypothetical prote
898	23	74.2	126	2	F70739	hypothetical prote
899	23	74.2	126	2	H90097	hypothetical prote
900	23	74.2	127	2	T35336	probable isomerase
901	23	74.2	128	2	T44877	hypothetical prote
902	23	74.2	129	2	F72435	2-amino-4-hydroxy-
903	23	74.2	132	2	B89834	teichoic acid bios
904	23	74.2	135	2	JC7671	acidicn cytochrom
905	23	74.2	135	2	CB3471	hypothetical prote

979	23	74.2	211	2	E75A15	hypothetical prote
980	23	74.2	211	2	F72239	hypothetical prote
981	23	74.2	212	2	T30700	hypothetical prote
982	23	74.2	212	2	T34052	hypothetical prote
983	23	74.2	213	2	S48402	hypothetical prote
984	23	74.2	213	2	T22991	hypothetical prote
985	23	74.2	215	2	C37390	transfer protein T
986	23	74.2	215	2	F96911	precortin-2 dehydr
987	23	74.2	216	1	C64081	L-fuculose phospho
988	23	74.2	218	1	F72214	hypothetical prote
989	23	74.2	221	2	S59075	splicing factor SR
990	23	74.2	222	2	AE3188	hypothetical prote
991	23	74.2	224	2	T10120	F420-dependent NAD
992	23	74.2	224	2	E72049	conserved hypother
993	23	74.2	224	2	F86575	CT691 hypothetical
994	23	74.2	226	2	T30615	hypothetical prote
995	23	74.2	230	2	E83594	probable pseudouri
996	23	74.2	231	2	G83809	MTA/SAH nucleosida
997	23	74.2	232	2	A69131	conserved hypother
998	23	74.2	232	2	S32963	hypothetical prote
999	23	74.2	233	2	T37376	27K structural pro
1000	23	74.2	234	2	T26363	hypothetical prote

ALIGNMENTS

RESULT 1

conserved hypothetical protein [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
 C:Accession: B90167
 R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Chan
 Joo, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R
 arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
 submitted to Genbank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: B90167
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <KUR>
 A:Cross-references: GB:AE006641; NID:g13813391; PIDN:AAK40593.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SS00254
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF0461

Query Match 100.0%; Score 31; DB 2; Length 281;
 Best Local Similarity 100.0%; Pred No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 160 VREEDL 165

RESULT 2

hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75453
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M. Shen, M.; Vamathevan, J.O.; Lam, P.; McDonald, L.; Uitterbeck, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75453
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-428 <WHI>

A:Cross-references: GB:AE001950; GB:AE000513; NID:g6458699; PIDN:AAF10548.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0969
 A:Map position: 1

Query Match 100.0%; Score 31; DB 2; Length 428;
 Best Local Similarity 100.0%; Pred No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 120 VREEDL 125

RESULT 3

N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
 C:Accession: J0261
 R:Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, O.W.; Mitsuo
 T. Biochem. 124, 670-678, 1998
 A:Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis
 A:Reference number: J0261; MUID:98391845; PMID:9722682
 A:Accession: J0261
 A:Molecule type: mRNA
 A:Residues: 1-484 <DCH>
 A:Cross-references: DDBJ:AB014679
 C:Comment: This protein catalyzes the transfer of sulfur of sulfate from 3'-phosphoad
 C:Superfamily: Chondroitin 6-sulfotransferase
 C:Keywords: sulfotransferase

Query Match 100.0%; Score 31; DB 2; Length 484;
 Best Local Similarity 100.0%; Pred No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 377 VREEDL 382

RESULT 4

conserved hypothetical protein HP0465 - Helicobacter pylori (strain 26895)
 C:Species: Helicobacter pylori
 C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C:Accession: A64578
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe
 son, U.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: A64578
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-631 <TOM>
 A:Cross-references: GB:AE000561; GB:AE000511; NID:g2113564; PIDN:AAD07527.1; PID:g231

Query Match 100.0%; Score 31; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred No. 54;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 182 VREEDL 187

RESULT 5

T30634
 hypothetical protein 32L - Molluscum contagiosum virus 1
 N:Alternate names: MC032L
 C:Species: Molluscum contagiosum virus 1
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jul-2000
 C:Accession: T30634
 R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
 A:Reference number: Z20876; MUID:96325459; PMID:8670425
 A:Accession: T30634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-748 <SEN>
 A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:NA055160.1; PID:g1491975
 C:Genetics:
 A:Note: MC032L
 C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 100.0%; Score 31; DB 2; Length 748;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 617 VREEDL 622

RESULT 6
 T21463
 hypothetical protein F28B1.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21463
 R:Matthews, L.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19426
 A:Accession: T21463
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-973 <WIL>
 A:Cross-references: EMBL:Z81517; PIDN:CA04209.1; GSPDB:GN00023; CESP:F28B1.2
 A:Experimental source: clone F28B1
 C:Genetics:
 A:Gene: CESP:F28B1.2
 A:Map position: 5
 A:Insertions: 70/3; 234/1; 267/3; 380/3; 600/3; 711/3; 926/3

Query Match 100.0%; Score 31; DB 2; Length 973;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 828 VREEDL 833

RESULT 7
 T30228
 polyketide synthase - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 C:Accession: T30228
 R:Paricio, J.F.; Molnar, I.; Scheweke, T.; Keenly, A.; Haydock, S.F.; Khaw, L.E.; Staun
 Gene 169, 9-16, 1996
 A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
 A:Reference number: Z20782; MUID:96186896; PMID:8635756
 A:Accession: T30228
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-6260 <APA>
 A:Cross-references: EMBL:X86780; NID:g987088; PID:g987102; PIDN:CAA60462.1
 C:Genetics:

A:Gene: rapC
 C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier pro
 C:Keywords: carrier protein
 F:53-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:1536-1607/Domain: acyl carrier protein homology <ACP1>
 F:1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:3123-3194/Domain: acyl carrier protein homology <ACP2>
 F:3238-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F:3728-3899/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F:5114-5185/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F:5242-5638/Domain: acyl carrier protein homology <ACP3>
 F:5759-6044/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F:6135-6206/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F:6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 100.0%; Score 31; DB 2; Length 6260;
 Best Local Similarity 100.0%; Pred. No. 6,7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 4365 VREEDL 4370

RESULT 8
 JC7628
 glucoside 3-dehydrogenase (EC 1.1.99.13) - Halomonas sp. alpha-15
 C:Species: Halomonas sp. alpha-15
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C:Accession: JC7628; PC7123
 R:Kojima, K.; Tsugawa, W.; Sode, K.
 Biochem. Biophys. Res. Commun. 282, 21-27, 2001
 A:Title: Cloning and expression of glucose 3-dehydrogenase from Halomonas sp. alpha-1
 A:Reference number: JC7628; MUID:21164693; PMID:11263965
 A:Accession: JC7628
 A:Molecule type: DNA
 A:Residues: 1-562 <KOJ>
 A:Accession: PC7123
 A:Molecule type: protein
 A:Residues: 2-11 <KO2>
 C:Comment: This enzyme, as an oligomeric enzyme composed of catalytic and electron tr
 and disaccharides.
 C:Genetics:
 A:Gene: q3dh
 C:Keywords: oxidoreductase

Query Match 96.8%; Score 30; DB 2; Length 562;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 142 VREEDL 147

RESULT 9
 HB7451
 oxidoreductase, GMC family CC1634 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: HB7451
 R:Nleman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko
 n, J.; Esmailaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: HB7451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-579 <STO>
 A:Cross-references: GB:AE005673; NID:g13423038; PIDN:AKK23612.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC1634

Query Match 96.8%; Score 30; DB 2; Length 579;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 153 IREYDL 158

RESULT 10
B71440
hypothetical protein d14605c - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: B71440
R:Byan, M.; Hancock, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, E.; Wedler, E.; Wambutt, R.; Wetzinger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalmatis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: B71440
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1322 <ENV>
A:Cross-references: GB:297342; NID:92245031; PID:92245072
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Arabidopsis thaliana hypothetical protein d14605c

Query Match 96.8%; Score 30; DB 2; Length 1322;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 21 IREYDL 26

RESULT 11
H87385
hypothetical protein CC1100 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87385
R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Lapid, M.I.; DeBoy, R.T.; Dodson, R.O.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87385
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-734 <STO>
A:Cross-references: GB:AE005673; NID:913422406; PID:AAK23084.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1100

Query Match 93.5%; Score 29; DB 2; Length 234;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 146 VREYDL 151

RESULT 12
F96020
probable iron ABC transporter periplasmic solute-binding protein precursor SMD20723 [
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F96020
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N₂-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: F96020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-355 <XU>
A:Cross-references: GB:AL591985; PID:CAQ49830.1; PID:915141318; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Punter, A.; Abola, P.; Ampe, F.; Barloy-Rub
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kries, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD20723
A:Genome: plasmid

Query Match 93.5%; Score 29; DB 2; Length 355;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 63 VREYDM 68

RESULT 13
T08203
polygalacturonase (EC 3.2.1.15) precursor - muskmelon
C:Species: Cucumis melo (muskmelon)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-May-2000
C:Accession: T08203
R:Hedfield, K.A.; Rose, J.K.; Yaver, D.S.; Berka, R.M.; Bennett, A.B.
Plant Physiol. 117, 363-373, 1998
A:Title: Polygalacturonase gene expression in ripe melon fruit supports a role for po
A:Reference number: 216403; MUID:98289082; PMID:9525689
A:Accession: T08203
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-393 <HAD>
A:Cross-references: EMBL:AF062465; NID:93320457; PID:93320458
C:Genetics:
A:Note: MP1
C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase

Query Match 93.5%; Score 29; DB 2; Length 393;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 329 VREYDI 334

RESULT 14
D84871
probable polygalacturonase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84871

R./lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Morfat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: D84871

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1405 <STC>

A:Cross-references: GB:AE002093; NID:g2281084; PIDN:AAB64020.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g43860

A:Map position: 2

C:Superfamily: polygalacturonase

Query Match 93.5%; Score 29; DB 2; Length 405;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 335 VREYDI 340

RESULT 15

G64446

ATP-dependent 26S proteasome regulatory subunit 4 homolog - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 19-Jan-2001

C:Accession: G64446

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.A>Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: G64446

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1430 <BUL>

A:Cross-references: GB:U67559; GB:L77117; NID:g1591798; PIDN:AAB99179.1; PID:g1591803; T

C:Genetics:

A:Map position: REV1116282-1114990

C:Superfamily: ATP-dependent 26S proteasome; FtsH/SEC18/CDC48-type ATP-binding domain hc

C:Keywords: ATP; nucleotide binding; P-loop

C:Keywords: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAMP>

F:211-218/Region: nucleotide-binding motif A (P-loop)

Query Match 93.5%; Score 29; DB 2; Length 430;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 167 VREYDI 172

RESULT 16

T39225

MAP kinase kinase kinase - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000

C:Accession: T39225

R:Churcher, C.M.; Gentile, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, August 1997

A:Reference number: 221837

A:Accession: T39225

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11401 <CHU>

A:Cross-references: EMBL:298763; PIDN:CAB11500.1; GSPDB:GN00066; SPDB:SPAC9G1.02

A:Experimental source: strain 972h; cosmid c9G1

C:Genetics:

A:Gene: SPDB:SPAC9G1.02

A:Map position: 1

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

Query Match 93.5%; Score 29; DB 2; Length 1401;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 1194 VREYDI 1199

RESULT 17

D71350

probable DNA-directed RNA polymerase, beta' subunit - *Syphilis spirochete*

C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)

C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 18-Jun-1999

C:Accession: D71350

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Ison, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; G they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998

A>Title: Complete genome sequence of *Treponema pallidum*, the *syphilis* spirochete.

A:Reference number: A71250; MUID:9832770; PMID:9665876

A:Accession: D71350

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1416 <COL>

A:Cross-references: GB:AE001205; GB:AE000520; NID:g3322501; PIDN:AAC65230.1; PID:g332

A:Experimental source: strain Nichols

A:Gene: TP0242

C:Superfamily: *Escherichia coli* DNA-directed RNA polymerase beta' chain

Query Match 93.5%; Score 29; DB 2; Length 1416;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 1052 VREYDI 1057

RESULT 18

B69441

conserved hypothetical protein AF1531 - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: B69441

R:Kleck, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod . Fleischmann, R.D.; Queckenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A:Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.

A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343; PMID:9385475

A:Accession: B69441

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1195 <KLE>

A:Cross-references: GB:AE000997; GB:AE000782; NID:92689320; PIDN:AAB89717.1; PID:g264

C:Superfamily: *Methanococcus jannaschii* conserved hypothetical protein MJ0038

Query Match 90.3%; Score 28; DB 1; Length 195;
Best Local Similarity 83.3%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||

Db 87 LRYEDL 92

RESULT 19

C:Species: *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: C66849

R:Botolin, A.; Winkler, P.; Mauger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ss

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: C66849

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-233 <STO>

A:Cross-references: GB:AE005176; PID:g12724820; PIDN:AAK05893.1; GSPDB:GN00146

A:Experimental source: strain IL1403

A:Gene: ysfD

Query Match 90.3%; Score 28; DB 2; Length 233;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 146 IRYEDI 151

RESULT 20

T2928 Immediate-early protein - ateline herpesvirus 3 (strain 73)

C:Species: ateline herpesvirus 3

A:Variety: strain 73

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T42928

R:Albrecht, J.C.; Fleckenstein, B.

Submitted to the EMBL Data Library, August 1998

A:Description: Primary structure of the herpesvirus ateles genome.

A:Reference number: 222274

A:Accession: T42928

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-273 <ALB>

A:Cross-references: EMBL:AF083424; PIDN:AA095536.1

A:Experimental source: strain 73

C:Superfamily: saimiri herpesvirus immediate-early protein 1

Query Match 90.3%; Score 28; DB 2; Length 273;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 208 VRYODL 213

RESULT 21

C69608 chitosanase csn - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: C69608

R:Kunst, F.; Ogatawa, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Acevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallet

iech, J.; Hartwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Roeha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tempere, P.; Tognoni, A.; Tostato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbel, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: C69608

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-277 <KUN>

A:Cross-references: GB:Z99117; GB:AL009126; NID:52634966; PIDN:CB14630.1; PID:ell839

A:Experimental source: strain 168

C:Genetics:

A:Gene: csn

Query Match 90.3%; Score 28; DB 2; Length 277;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 222 VRYDDL 227

RESULT 22

A83634 hypothetical protein PA0100 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83634

R:Scovier, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83634

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-306 <STO>

A:Cross-references: GB:AE004448; GB:AE004091; NID:g9945917; PIDN:AA03490.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0100

Query Match 90.3%; Score 28; DB 2; Length 306;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 16 VRYODL 21

RESULT 23

A84523 probable steroid sulfotransferase [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: A84523

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujita, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:1061157

A:Accession: A84523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: GB:AE002093; NID:g3650034; PIDN:AA061289.1; GSPDB:GN00139

C:Genetics:
A:Gene: At2g14920
A:Map position: 2
C:Superfamily: alcohol sulfotransferase

Query Match 90.3%; Score 28; DB 2; Length 333;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 232 IRYEDL 237

RESULT 24

C38453
anaerobic sulfite reduction protein C - *Salmonella typhimurium*
C:Species: *Salmonella typhimurium*
C>Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 17-Mar-2000
C:Accession: C38453

R: Huang, C.J.; Barrett, E.L.
J. Bacteriol. 173, 1544-1553, 1991
A:Title: Sequence analysis and expression of the *Salmonella typhimurium* asr operon encod
A:Reference number: A38453; MUID: 91139599; PMID: 1704886

A:Accession: C38453
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <HUA>
A:Cross-references: GB:M57706; NID:g153881; PIDN:AA9277.1; PID:g153884

C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein
F:173-230/Domain: ferredoxin 2[4Fe-4S] homology <FER>
F:180,183,186,222/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F:190,212,215,218/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 90.3%; Score 28; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 69 IRYEDL 74

RESULT 25

AE0825
anaerobic sulfite reductase chain C (EC 1.8.1.-) [imported] - *Salmonella enterica* subsp.
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0825

R: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, R.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moulé, S.; O'Garra, P.
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; PMID: 11677608
A:Accession: AE0825

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02753.1; PID:g16503764; GSPDB:GN00176

C:Genetics:
A:Gene: asrC
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
C:Keywords: oxidoreductase

Query Match 90.3%; Score 28; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

DB 69 IRYEDL 74

RESULT 26

T29003
hypothetical protein ZC513.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29003

R: Wu, X.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid ZC513.
A:Reference number: Z20551

A:Accession: T29003
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-346 <WUX>
A:Cross-references: EMBL:U53155; PIDN:AA48271.1; GSPDB:GN00023; CESP:ZC513.9

C:Genetics:
A:Gene: CESP:ZC513.9
A:Map position: 5
A:Introns: 62/3; 117/2; 252/3; 312/1

Query Match 90.3%; Score 28; DB 2; Length 346;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 79 IRYEDL 84

RESULT 27

G70729
hypothetical protein RV2267c - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: G70729

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; MUID: 98295987; PMID: 9634230
A:Accession: G70729

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-388 <COL>
A:Cross-references: GB:Z77163; GB:AL123456; NID:g3261610; PIDN:CAB00968.1; PID:e25507

C:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2267c

Query Match 90.3%; Score 28; DB 2; Length 386;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 302 IRYEDL 307

RESULT 28

AH2447
molybdopterin biosynthesis protein [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH2447

R: Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpco, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 6, 205-213, 2001
A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anaerostipes*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-436 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA576835.1; PID:g17134274; GSPDB:GN00179
A:Experimental source: strain FCC 7120
C:Genetics: moea
C:Superfamily: molybdenum cofactor biosynthesis protein moea-2

Query Match 90.3%; Score 28; DB 2; Length 456;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
DB 84 VREYDV 89

RESULT 29

C75180
tldd related protein PAB1955 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75180
R:anonymus; Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: C75180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <KAM>
A:Cross-references: GB:AJ248284; GB:AL096836; NID:95457730; PIDN:CB49522.1; PID:g545803
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1955
C:Superfamily: Escherichia coli tldd protein

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
DB 25 LREYDL 30

RESULT 30

A57397
chondroitin 6-sulfotransferase (EC 2.8.2.17) precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 20-Jun-2000
C:Accession: A57397
R:Fukuta, M.; Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuchi,
J. Biol. Chem. 270, 18575-18580, 1995
A:Title: Molecular cloning and expression of chick chondrocyte chondroitin 6-sulfotransf
A:Reference number: A57397; MUID:95355490; PMID:7629189
A:Accession: A57397
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-458 <FUK>
A:Cross-references: GB:D49915; NID:9711262; PIDN:BA08655.1; PID:g9711263
C:Superfamily: chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
DB 348 VREYDV 353

RESULT 31

JC7350
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: JC7350
R:Uchimura, K.; Fasakhany, F.; Kadamatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa,
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of
A:Reference number: JC7350; MUID:20374462; PMID:10913333
A:Accession: JC7350
A:Molecule type: mRNA
A:Residues: 1-484 <UCH>
A:Cross-references: DDBJ:AB040710
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosyl
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 90.3%; Score 28; DB 2; Length 484;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
DB 371 LREYDL 376

RESULT 32

JC7351
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
C:Species: Homo sapiens (man)
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: JC7351
R:Uchimura, K.; Fasakhany, F.; Kadamatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa,
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of
A:Reference number: JC7350; MUID:20374462; PMID:10913333
A:Accession: JC7351
A:Molecule type: mRNA
A:Residues: 1-486 <UCH>
A:Cross-references: DDBJ:AB040711
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosyl
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 90.3%; Score 28; DB 2; Length 486;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
DB 373 LREYDL 378

RESULT 33

F90169
hypothetical protein tgra [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90169
R:Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90169

Query Match 90.3%; Score 28; DB 2; Length 486;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <RUR>
A:Cross-references: GB:AE006641; NID:g13813415; PIDN:AAK40613.1; GSPDB:GN00155
C:Genetics:
A:Gene: tga

Query Match 90.3%; Score 28; DB 2; Length 503;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
:|||||
DB 6 VKYEDL 11

RESULT 34

E89804
alkyl hydroperoxide reductase subunit F [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence-revision 10-May-2001 #text-change 22-Oct-2001
C:Accession: E89804
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89756; MUID:21311952; PMID:11418146
A:Accession: E89804
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <KUR>
A:Cross-references: GB:BA000018; PID:g13700294; PIDN:BA41592.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: amp
C:Superfamily: NADH oxidase (hydrogen peroxide-forming); thioredoxin reductase homology

Query Match 90.3%; Score 28; DB 2; Length 507;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
:|||||
DB 416 IREEDM 421

RESULT 35

T45641
beta-D-glucan exohydrolase-like protein - Arabidopsis thaliana
N:Alternate names: protein F13112.100
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence-revision 04-Feb-2000 #text-change 21-Jul-2000
C:Accession: T45641
R:Choinsne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cottolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z23010
A:Accession: T45641
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <CHO>
A:Cross-references: EMBL:AL133292
A:Experimental source: cultivar Columbia; BAC clone F13112
C:Genetics:
A:Map position: 3
A:Intons: 56/1; 124/2; 154/3; 203/2; 284/3; 319/3; 380/3; 446/1
A:Note: F13112.100
C:Superfamily: beta-glucosidase

Query Match 90.3%; Score 28; DB 2; Length 612;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
:|||||
DB 230 LREEDL 235

RESULT 36

B91271
probable periplasmic binding protein [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence-revision 18-Jul-2001 #text-change 18-Jul-2001
C:Accession: B91271
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B91271
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1107 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA38561.1; PID:g13364615; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ecs5138

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
:|||||
DB 348 LREEDL 353

RESULT 37

B86112
probable periplasmic binding protein yJep [imported] - Escherichia coli (strain O157:
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 14-Sep-2001
C:Accession: B86112
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallant, E.; Potamousts, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B86112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1107 <STO>
A:Cross-references: GB:AE005174; NID:g12519144; PIDN:AG59358.1; GSPDB:GN00145; UMG:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yJep

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
:|||||
DB 348 LREEDL 353

RESULT 38

E65226
hypothetical 123.8 kD protein in genx-psd intergenic region - Escherichia coli (strai
N:Alternate names: hypothetical protein F1107
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence-revision 17-Sep-1997 #text-change 01-Mar-2002
C:Accession: E65226; S56387
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E65226
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1107 <BLAT>
A:Cross-references: GB:AE000489; GB:U00096; NID:g2367354; PIDN:AAC77119.1; PID:g2367355
A:Experimental source: strain K-12, substrain MG1655
A:Burkhardt, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.
A:Reference number: S56314; MUID:95333362; PMID:7610040
A:Accession: S56387
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-746, 'x', 748-1014, 'A', 1016-1107 <BUR>
A:Cross-references: EMBL:U14003; NID:g1263172; PIDN:AA97058.1; PID:g537003
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C:Genetics:
A:Gene: yJep
A:Start codon: GTG

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 348 IREEDL 353

RESULT 39
ZIN2P3
genome polypeptide - parainfluenza virus type 3
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: parainfluenza virus type 3
A:Note: host Homo sapiens (man)
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 29-May-1998
C:Accession: A29246
R:Galinski, M.S.; Mink, M.A.; Pons, M.W.
Virology 165, 499-510, 1988
A:Title: Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene
A:Reference number: A29246; MUID:88306242; PMID:2841798
A:Accession: A29246
A:Molecule type: genomic RNA
A:Residues: 1-2233 <GML>
A:Cross-references: GB:M21649
C:Genetics:
A:Gene: L
C:Superfamily: parainfluenza virus RNA-directed RNA polymerase
C:Keywords: ATP; nucleotidyltransferase

Query Match 90.3%; Score 28; DB 1; Length 2233;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 1126 IREEDL 1131

RESULT 40
H83960
hypothetical protein BH2488 [Imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H83960
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20513582; PMID:11058132
A:Accession: H83960
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-73 <STO>
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06207.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2488

Query Match 87.1%; Score 27; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IREEDL 6
DB 39 IREEDL 43

RESULT 41
F72649
hypothetical protein APE0628 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: F72649
R:Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72649
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BA79598.1; PID:d1043384; PID:g
C:Experimental source: strain K1
C:Genetics:
A:Gene: APE0628
C:Superfamily: *Aeropyrum pernix* hypothetical protein APE0628

Query Match 87.1%; Score 27; DB 2; Length 108;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 16 IREEDL 21

RESULT 42
H69433
response regulator homolog - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Aug-1999
C:Accession: H69433
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
gloek, A.; Zhou, L.; Overbeck, R.; Gocayne, J.D.; Weisman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ueberlack, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69433
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-133 <KLE>
A:Cross-references: GB:AE001001; GB:AE000782; NID:g2689324; PIDN:AAB9775.1; PID:g264
C:Superfamily: chemotaxis chey protein; response regulator homolog
C:Keywords: phosphoprotein
F:4.11/Domain: response regulator homolog <RR>
F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 87.1%; Score 27; DB 2; Length 133;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

A:Accession: G70340
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-225 <AOF>
A:Cross-references: GB:AE000690; NID:92983100; PIDN:AAC06713.1; PID:92983108; GB:AE00065
A:Experimental source: strain VFS
A:Genetics:
A:Gene: mgcC
C:Superfamily: Mg2+-transporting ATPase

Query Match 87.1%; Score 27; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRYED 6
|||||
Db 202 VRYED 206

RESULT 48
S28679
glycoprotein 30 - phage SP01
C:Species: phage SP01
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
C:Accession: S28679
R:Scarlato, V.; Sayre, M.H.
Gene 114, 115-119, 1992
A:Title: Sequence of the bacteriophage SP01 gene 30.
A:Reference number: S28678; MUID:92267370; PMID:1587473
A:Accession: S28679
A:Molecule type: DNA
A:Residues: 1-225 <SCA>
A:Cross-references: EMBL:M82842; NID:9216115; PIDN:AAA32597.1; PID:9216117
C:Genetics:
A:Start codon: TTG
C:Superfamily: phage SP01 glycoprotein 30
C:Keywords: glycoprotein

Query Match 87.1%; Score 27; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 32 VRYED 36

RESULT 49
C90026
hypothetical protein narQ [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C90026
R:Kurita, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C90026
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <KUR>
A:Cross-references: GB:BA000018; PID:913702080; PIDN:BA843372.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: narQ
C:Superfamily: Escherichia coli fdhD protein

Query Match 87.1%; Score 27; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 12 VRYED 16

RESULT 50
S07014
hypothetical protein 273 (orf1 5' region) - Escherichia coli plasmid F
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jul-2000
C:Accession: S07014
R:Loeb, S.; Cram, D.; Skurray, R.
Mol. Gen. Genet. 219, 177-186, 1989
A:Title: Nucleotide sequence of the leading region adjacent to the origin of transfer
A:Reference number: S07013; MUID:90136505; PMID:2693941
A:Accession: S07014
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-273 <LOH>
A:Cross-references: GB:AF106329; GB:M97768; GB:M21120; GB:X17189; NID:95702165; PIDN:
A:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia hypothetical protein yafZ

Query Match 87.1%; Score 27; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 189 VRYED 193

Search completed: February 20, 2003, 13:34:39
Job time : 26.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:36 ; Search time 2.57143 Seconds
(without alignments)
96.778 Million cell updates/sec

```
Title: US-09-816-825-9
Perfect score: 31
Sequence: 1 VRYEDL 6
```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 1000 summaries
```

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	29	93.5	430	1	PSMR_MERJA	O58576 methanococc
2	29	93.5	1401	1	WIS4_SCHDO	O14299 schizosacch
3	29	93.5	1416	1	RPOC_TREPA	O08327 treponema f
4	28	90.3	277	1	CHIS_BASCU	O07921 bacillus su
5	28	90.3	306	1	CAC3_STPNV	O54800 streptococc
6	28	90.3	337	1	ASRC_SALTU	O26476 salmonella
7	28	90.3	388	1	YMG7_MCTU	O50655 mycobacteri
8	28	90.3	436	1	MOEA_ANASP	O44243 anabena str
9	28	90.3	458	1	C6ST_CHICK	O92179 gallus gall
10	28	90.3	507	1	AHPF_STAPU	O05204 staphylococ
11	28	90.3	626	1	MAK3_MOUSE	O61084 mus musculi
12	28	90.3	1107	1	YUEP_ECOLI	P59285 escherichia
13	28	90.3	2233	1	RRPL_P13H4	P12577 human para
14	27	87.1	157	1	URPE_CORGL	O97hm3 corynebact
15	27	87.1	159	1	MLE_TODPA	O50595 toderodes f
16	27	87.1	225	1	GP30_BPSP1	O38423 bacterioph
17	27	87.1	273	1	YUEB_ECOLI	P18005 escherichia
18	27	87.1	308	1	GDFE_HUMAN	O93988 homo sapien
19	27	87.1	311	1	MRBW_BACSU	O007876 bacillus s
20	27	87.1	341	1	REBB_NEITMA	O95662 neisseria n
21	27	87.1	346	1	REBB_NEIGO	P37612 neisseria n
22	27	87.1	355	1	REBB_NEITMB	P55224 neisseria n
23	27	87.1	396	1	PGK_STAM	O92584 staphylococ
24	27	87.1	418	1	LE21_ARFTU	O28316 archaeglo
25	27	87.1	418	1	LE21_MERKA	O8Lvt2 methanopyr
26	27	87.1	428	1	LE22_MERTH	O21668 methanobact
27	27	87.1	439	1	DCDA_BACHD	O96cm5 bacillus ha
28	27	87.1	439	1	GUAD_ECOLI	P76641 escherichia
29	27	87.1	492	1	C133_DROME	O9vwp3 streptococ
30	27	87.1	493	1	FUT3_ARATH	O9c371 arabidopsis
31	27	87.1	514	1	Y213_HAIAI	P44572 haemophilus
32	27	87.1	621	1	PPCK_PYRHO	O58050 pyrococcus
33	27	87.1	623	1	SKNI_CAEBL	P34707 caenorhabdi

107	25	80.6	346	1	YQCL_BACSU	P45946	bacillus su	180	24	77.4	235	1	VG53_BPMD2	O64243	mycobacteri
108	25	80.6	350	1	Y535_MENUA	Q57975	methanococc	181	24	77.4	245	1	CRS3_HUMAN	P54108	homo sapien
109	25	80.6	366	1	YVAF_BACSU	P37518	bacillus su	182	24	77.4	245	1	PSBP_CHURE	P14711	chlamydomon
110	25	80.6	375	1	PCF_TACTR	P21902	tachypleus	183	24	77.4	257	1	NUDC_BCO57	Q8x6x7	escherichia
111	25	80.6	409	1	EFTU_EUGGR	P02991	euglena gra	184	24	77.4	257	1	NUDC_ECOLI	P32664	escherichia
112	25	80.6	414	1	PYRC_TREAC	Q9hmo0	thermoplasm	185	24	77.4	257	1	NUDC_SALTI	Q8x328	salmonella
113	25	80.6	436	1	THCB_RHOER	P43492	rhodococcus	186	24	77.4	257	1	NUDC_SALTY	Q91915	salmonella
114	25	80.6	452	1	SELA_AQUAE	O67140	rhodococcus	187	24	77.4	257	1	ELAH_ECOLI	P43530	escherichia
115	25	80.6	486	1	DNAB_HELPJ	Q92jms	helicobacte	188	24	77.4	258	1	ELAP_ECOLI	P06717	escherichia
116	25	80.6	488	1	DNAB_HELPY	O25916	helicobacte	189	24	77.4	258	1	PSBP_LYCES	P26795	lycopersico
117	25	80.6	491	1	GLNA_ARCFU	O29313	archaeoglob	190	24	77.4	258	1	PSBP_WHEAT	Q00434	tritlicum ae
118	25	80.6	520	1	GAL3_YEAST	P13045	saccharomyc	191	24	77.4	259	1	PSBP_PEA	P16059	pisum sativ
119	25	80.6	532	1	YMO0_YEAST	O04458	saccharomyc	192	24	77.4	259	1	Y128_CHLPN	P11594	chlamydia p
120	25	80.6	547	1	TEPH_MOUSE	Q9dbag	homo musculu	193	24	77.4	260	1	PSBP_SINAL	P13594	sinapis alb
121	25	80.6	548	1	TEPH_MOUSE	P32780	homo sapien	194	24	77.4	260	1	PSBP_SALTU	P93566	solanum tub
122	25	80.6	573	1	GIP1_YEAST	P38229	saccharomyc	195	24	77.4	263	1	PSBP_CUGSA	O9s198	cucumis sat
123	25	80.6	585	1	YHD9_YEAST	P37332	saccharomyc	196	24	77.4	264	1	PSBP_PRIAG	O45080	trilliarita
124	25	80.6	605	1	VE1_HPV09	Q05111	human papil	197	24	77.4	265	1	PSBP_NARPS	Q04047	narcissus p
125	25	80.6	609	1	VE1_HPV17	O02512	human papil	198	24	77.4	265	1	SANT_PLAFR	P04927	plasmidum
126	25	80.6	616	1	ORC3_YEAST	P54750	saccharomyc	199	24	77.4	267	1	PSBP_SPIOL	P12302	spiniacia ol
127	25	80.6	620	1	YGI8_YEAST	P53207	saccharomyc	200	24	77.4	268	1	YBGF_PSEPU	P43037	pseudomonas
128	25	80.6	626	1	MK33_HUMAN	Q99729	saccharomyc	201	24	77.4	271	1	KKAI_ECOLI	P00551	escherichia
129	25	80.6	633	1	CNE7_HUMAN	Q9ud16	homo sapien	202	24	77.4	271	1	KKAI_SALTY	Q03447	salmonella
130	25	80.6	716	1	SYEC_SCHPO	O13775	schizosacch	203	24	77.4	272	1	KKAB_ECOLI	P14509	escherichia
131	25	80.6	727	1	NDAM_BOVIN	P15690	bos taurus	204	24	77.4	271	1	T2F1_HAEIN	P20588	haemophilus
132	25	80.6	883	1	RPOL_BPT7	P00573	bacterioph	205	24	77.4	280	1	TRPA_CORGL	P06562	corynebacte
133	25	80.6	884	1	RPO6_YEAST	P07659	bacterioph	206	24	77.4	288	1	LPXC_RICCN	O921t1	ricketsia
134	25	80.6	928	1	PR03_YEAST	P19735	saccharomyc	207	24	77.4	291	1	MSAB_TREPA	O83641	t.peptide m
135	25	80.6	928	1	PR03_YEAST	P40048	saccharomyc	208	24	77.4	293	1	YMO2_RHIME	O83641	t.peptide m
136	25	80.6	979	1	PL15_MYCHR	P41508	mycoplasma	209	24	77.4	296	1	ILVR_CAUCR	P52670	caulobacter
137	25	80.6	988	1	SY1_SYNY3	P73505	synchocyst	210	24	77.4	300	1	RIP3_MAIZE	P25891	zea mays (m
138	25	80.6	1013	1	YTI93_DEIRA	Q9r1y9	deinococcus	211	24	77.4	301	1	RIPX_MAIZE	P25892	zea mays (m
139	25	80.6	1095	1	CC25_SACKU	O02342	saccharomyc	212	24	77.4	303	1	LPXC_PSEAE	P47200	pseudomonas
140	25	80.6	1249	1	TPP2_HUMAN	P29144	homo sapien	213	24	77.4	304	1	RIP9_MAIZE	P25892	zea mays (m
141	25	80.6	1249	1	TPP2_RAT	O64560	rattus norv	214	24	77.4	306	1	LPXC_SHEVI	O96v47	shevanelia
142	25	80.6	1262	1	TPP2_MOUSE	O64514	mus musculu	215	24	77.4	306	1	RDGC_PSEAE	O96v47	shevanelia
143	25	80.6	1302	1	UBPY_CAEEL	O09931	caenorhabdi	216	24	77.4	309	1	SANT_PLAFN	P04928	plasmidum
144	25	80.6	1331	1	XDH_BOVIN	R80457	bos taurus	217	24	77.4	321	1	VF05_YACCC	P21014	vaccinia vi
145	25	80.6	1537	1	DNM1_CHICK	O92072	gallus gall	218	24	77.4	321	1	VF05_YACCP	O00320	vaccinia vi
146	25	80.6	1690	1	RPOC_THEMA	P36252	thermotoga	219	24	77.4	322	1	VF05_YACCV	P24358	vaccinia vi
147	25	80.6	1748	1	YNR2_YEAST	P53866	saccharomyc	220	24	77.4	322	1	VP05_YARV	P33865	variola vir
148	25	80.6	1803	1	YTI3_YEAST	P47074	saccharomyc	221	24	77.4	326	1	HEM2_PASMT	P57874	pasteurella
149	24	77.4	83	1	YB62_YEAST	O58857	methanococc	222	24	77.4	333	1	HYU4_LACHE	P22294	lactobacilli
150	24	77.4	102	1	RPOL_AERPE	Q9yea1	aeropyrum p	223	24	77.4	342	1	YEA4_YEAST	P40004	saccharomyc
151	24	77.4	110	1	VCAD_LAMB	P03712	bacterioph	224	24	77.4	343	1	YNA1_ECOLI	P77223	escherichia
152	24	77.4	127	1	PAND_ANASP	O8yr79	anabaena sp	225	24	77.4	355	1	BUR1_CIOAB	Q45829	clostridium
153	24	77.4	147	1	RUL3_STRCO	O53874	streptomyce	226	24	77.4	359	1	MSAB_HELPJ	O92mk8	h.peptide m
154	24	77.4	153	1	MLE_CAEEL	P53014	caenorhabdi	227	24	77.4	359	1	MSAB_HELPY	O25011	h.peptide m
155	24	77.4	156	1	MLE_AEOTR	P07291	aequipecten	228	24	77.4	360	1	VIS_AGRVT	O04554	agrobacteri
156	24	77.4	156	1	MLE_PATYE	P07280	palinopecten	229	24	77.4	374	1	HS83_DROVI	P04811	drosophila
157	24	77.4	171	1	NIDM_HUMAN	O96000	homo sapien	230	24	77.4	375	1	HS83_DROPS	P04809	drosophila
158	24	77.4	177	1	AHPD_MYCTU	O57353	mycobacteri	231	24	77.4	375	1	HS83_DROSI	P04810	drosophila
159	24	77.4	185	1	THE2_ARATH	O9xfn9	arabidopsi	232	24	77.4	375	1	SANT_PLAFV	P09593	plasmidum
160	24	77.4	188	1	Y084_NPVAC	P41474	autographa	233	24	77.4	390	1	MEIL_HUMAN	O00470	homo sapien
161	24	77.4	190	1	NDGM_ARATH	O95748	arabidopsi	234	24	77.4	390	1	MEIL_MOUSE	O60954	mus musculu
162	24	77.4	190	1	NDGM_ORISA	Q35322	oryza sativ	235	24	77.4	390	1	MEIL_XENLA	P75937	xenopus lae
163	24	77.4	190	1	NDGM_SOLTU	R80261	solanum tub	236	24	77.4	391	1	MEIK_RHITO	O96a80	rhizobium l
164	24	77.4	192	1	NDGM_BETTR	Q33994	beta tritigyn	237	24	77.4	395	1	NDK7_MOUSE	O96x18	mus musculu
165	24	77.4	192	1	NDGM_BETVU	Q37787	beta vulgar	238	24	77.4	395	1	NDK7_RAT	O96x17	rattus norv
166	24	77.4	192	1	NDGM_BETWE	O34011	beta webbia	239	24	77.4	398	1	DXR_ECO57	O88xyl	escherichia
167	24	77.4	192	1	NDGM_PROMI	Q37622	protitheca	240	24	77.4	398	1	DXR_ECOLI	P45568	escherichia
168	24	77.4	192	1	SCPA_PENSP	P02636	penaeus sp.	241	24	77.4	398	1	DXR_SALTI	O8x9a6	salmonella
169	24	77.4	193	1	YAJB_ECOLI	P21515	escherichia	242	24	77.4	398	1	DXR_SALTY	O8zrp3	salmonella
170	24	77.4	195	1	NDGM_MARPO	P34944	marichantia	243	24	77.4	398	1	PKR_STRPN	O97889	streptococc
171	24	77.4	199	1	TPM1_YEAST	P17536	saccharomyc	244	24	77.4	399	1	FAST_TROFA	P46373	rhodococcus
172	24	77.4	204	1	NDGM_RECAM	O21271	reclinomona	245	24	77.4	400	1	TRA6_BACST	Q45618	bacillus st
173	24	77.4	212	1	MSRA_SALTI	O82150	salmonella	246	24	77.4	405	1	CPYK_SACER	P33371	saccharopol
174	24	77.4	213	1	ID1_HALNI	O9h540	halobacteri	247	24	77.4	405	1	YLF2_YEAST	P33746	saccharomyc
175	24	77.4	215	1	FUCA_ECOLI	P19560	escherichia	248	24	77.4	409	1	Y291_METUA	O57733	methanococc
176	24	77.4	217	1	PSBP_BRAJU	Q96334	brassica ju	249	24	77.4	410	1	SAHH_TREAC	O9hxx4	thermoplasm
177	24	77.4	228	1	YB62_CAEEL	Q21683	caenorhabdi	250	24	77.4	410	1	SAHH_TREVO	O97924	thermoplasm
178	24	77.4	232	1	YF83_AQUAE	O67008	aquifex aeo	251	24	77.4	417	1	SYS_MYCCE	P47251	mycoplasma
179	24	77.4	234	1	TNF8_HUMAN	P32971	homo sapien	252	24	77.4	424	1	VG46_BPPF1	P25131	bacterioph

253	24	77.4	425	1	PURA_FUSNN	P58793	fusobacteri	326	24	77.4	723	1	HS9B_RAT	P34058	rattus norv
254	24	77.4	425	1	PURA_FUSND	O65881	fusobacteri	327	24	77.4	725	1	HS9B_BRARE	O57521	brachydanio
255	24	77.4	428	1	DHE4_SYNX3	P54386	synechocyst	328	24	77.4	725	1	HS9B_CHICK	O04619	gallus gall
256	24	77.4	432	1	PURA_YEAST	P80210	saccharomyc	329	24	77.4	726	1	HS9A_BRARE	O00474	brachydanio
257	24	77.4	443	1	MCRB_METVA	P07936	methanococc	330	24	77.4	728	1	HS9A_CHICK	P11501	gallus gall
258	24	77.4	443	1	MCRB_METVO	P11561	methanococc	331	24	77.4	732	1	ACET_HUMAN	P22966	homo sapien
259	24	77.4	443	1	PC17_HUMAN	Q01850	homo sapien	332	24	77.4	732	1	HS9A_CRIGR	P66633	cricetulus
260	24	77.4	451	1	CLUS_COTJA	P14018	coturnix co	333	24	77.4	732	1	HS9A_MOUSE	P07901	mus musculu
261	24	77.4	458	1	CHPN_HUMAN	P15169	homo sapien	334	24	77.4	732	1	HS9A_PIG	O02705	sus scrofa
262	24	77.4	458	1	NMPB_ENTHR	O08637	enterococcu	335	24	77.4	735	1	MCMS_YEAST	P29468	saccharomyc
263	24	77.4	464	1	RCCL_CANAL	P52499	candida alb	336	24	77.4	776	1	CG2F_MOISE	P11944	mus musculu
264	24	77.4	465	1	HKK4_HUMAN	P35537	homo sapien	337	24	77.4	783	1	TRE2_THESC	Q30654	thermoplasm
265	24	77.4	465	1	HKK4_MOUSE	P52792	mus musculu	338	24	77.4	785	1	CAD7_CHICK	O07663	gallus gall
266	24	77.4	468	1	GINA_METCA	P15124	methylococc	339	24	77.4	785	1	CAD7_HUMAN	Q01458	homo sapien
267	24	77.4	470	1	YMPB_CAEEL	P53993	caenorhabdi	340	24	77.4	786	1	CG2F_HUMAN	P41003	homo sapien
268	24	77.4	471	1	MTHB_DROME	P83118	drosophila	341	24	77.4	789	1	K6PE_HABCO	O27665	haemophilu
269	24	77.4	472	1	ARPB_FERIS	O50341	fervidobact	342	24	77.4	790	1	YDDB_ECOLI	P31827	escherichia
270	24	77.4	472	1	CBIR_FELCA	O02777	fells silve	343	24	77.4	801	1	CADK_HUMAN	Q48740	homo sapien
271	24	77.4	472	1	CBIR_HUMAN	P12534	homo sapien	344	24	77.4	862	1	ADHE_CLOAB	O33744	clostridium
272	24	77.4	472	1	ELBL_ADE41	P10546	human adeno	345	24	77.4	901	1	GREB_BORBU	O51157	borrelia bu
273	24	77.4	473	1	CBIR_MOUSE	P47746	mus musculu	346	24	77.4	908	1	DPOL_BORBU	O51498	borrelia bu
274	24	77.4	473	1	CBIR_POEGU	P56971	poephila gu	347	24	77.4	946	1	ITRH_HUMAN	P33307	homo sapien
275	24	77.4	473	1	CBIR_RAT	P20272	rattus norv	348	24	77.4	960	1	CSEL_YEAST	P33307	saccharomyc
276	24	77.4	497	1	Y124_YEAST	P47114	saccharomyc	349	24	77.4	982	1	SUDJ_CABRL	P66549	caenorhabdi
277	24	77.4	508	1	NOG0_DROME	O44081	drosophila	350	24	77.4	1032	1	TLR8_MOUSE	P36862	mus musculu
278	24	77.4	508	1	YBCK_ECOLI	P73698	escherichia	351	24	77.4	1043	1	RRPO_NODAV	Q31mm4	nodamura vi
279	24	77.4	509	1	DKC1_RAT	P46615	rattus norv	352	24	77.4	1077	1	YBK1_YEAST	P18167	saccharomyc
280	24	77.4	514	1	DKC1_HUMAN	O60832	homo sapien	353	24	77.4	1082	1	SP23_YEAST	P52303	saccharomyc
281	24	77.4	515	1	PEN3_ADEG1	O64755	avian adeno	354	24	77.4	1097	1	RNT1_FUGRU	Q98tr3	fugu rubrip
282	24	77.4	517	1	UZAF_SCHPO	P36629	schizosacch	355	24	77.4	1108	1	UBPB_SCHPO	O09879	schizosacch
283	24	77.4	521	1	CEAL_SHISO	P21178	shigella so	356	24	77.4	1113	1	MG42_YEAST	P40578	saccharomyc
284	24	77.4	521	1	RE3_STRAU	O66490	staphylococ	357	24	77.4	1113	1	RNT1_MOUSE	Q9epu8	mus musculu
285	24	77.4	524	1	CP51_UNCNE	O14442	uncinula ne	358	24	77.4	1129	1	RNT1_HUMAN	Q92900	homo sapien
286	24	77.4	524	1	PPCK_CAME	O9PP01	campylobact	359	24	77.4	1145	1	Y623_DROME	Q67420	homo sapien
287	24	77.4	525	1	PEN3_ADEGX	P33538	avian adeno	360	24	77.4	1184	1	ALAC_ARATH	O03792	arabidopsis
288	24	77.4	531	1	UD16_RABIT	Q26611	oryctolagus	361	24	77.4	1195	1	ROB_THESC	O07587	thermoplasm
289	24	77.4	534	1	CN9A_MOUSE	O70628	mus musculu	362	24	77.4	1206	1	FM14_MOUSE	O05855	mus musculu
290	24	77.4	536	1	SYEM_YEAST	P48525	saccharomyc	363	24	77.4	1213	1	FMN_CHICK	O05855	gallus gall
291	24	77.4	542	1	Y159_METTA	Q57623	methanococc	364	24	77.4	1221	1	DBP2_KLULA	O42726	Kluveromyc
292	24	77.4	546	1	ALKK_PSEOL	Q00594	pseudomonas	365	24	77.4	1242	1	NEPN_MOUSE	Q9qz87	mus musculu
293	24	77.4	548	1	CH60_ZYMO	P48220	zymomonas m	366	24	77.4	1265	1	REOD_CYAPA	P81291	cyanophora
294	24	77.4	550	1	CRES_DICDI	P21837	dictyosteli	367	24	77.4	1279	1	BSC3_PSEPL	P58237	pseudomonas
295	24	77.4	588	1	RUBB_BRANA	P21241	brassica na	368	24	77.4	1306	1	ACE_HUMAN	P12821	homo sapien
296	24	77.4	589	1	IE2P_THESC	Q97b60	thermoplasm	369	24	77.4	1318	1	VIVD_BPT7	P23286	bacterioph
297	24	77.4	589	1	IE2P_THESC	Q97b60	thermoplasm	370	24	77.4	1330	1	XDH_RAT	P23286	rattus norv
298	24	77.4	589	1	IE2P_THESC	Q97b60	thermoplasm	371	24	77.4	1335	1	XDH_MOUSE	P23286	rattus norv
299	24	77.4	593	1	CN9A_HUMAN	P23731	ascaris suu	372	24	77.4	1390	1	N155_RAT	Q00519	mus musculu
300	24	77.4	593	1	SANT_PLAF7	O76083	homo sapien	373	24	77.4	1391	1	N155_HUMAN	P71199	rattus norv
301	24	77.4	595	1	RUBB_PEA	O06927	pisum sativ	374	24	77.4	1391	1	N155_MOUSE	O75694	homo sapien
302	24	77.4	597	1	TYPA_SYNX3	P72749	synechocyst	375	24	77.4	1451	1	SPT6_YEAST	P23615	saccharomyc
303	24	77.4	600	1	RUBB_ARATH	P21240	arabidopsis	376	24	77.4	1468	1	FMN1_MOUSE	O05866	mus musculu
304	24	77.4	607	1	LPXB_CHILM	Q97j44	chlamydia m	377	24	77.4	1679	1	Y109_YEAST	P40457	saccharomyc
305	24	77.4	613	1	DP2S_PTYRO	P81412	pyrococcus	378	24	77.4	1736	1	ZOI_HUMAN	O07157	homo sapien
306	24	77.4	619	1	PECK_PYRAB	Q9uy53	pyrococcus	379	24	77.4	1828	1	MY5A_RAT	Q07413	homo sapien
307	24	77.4	618	1	SEDB_HAEIN	P43927	haemophilus	380	24	77.4	1829	1	MY5A_CHICK	Q04640	gallus gall
308	24	77.4	631	1	TAC3_MOUSE	Q97j11	mus musculu	381	24	77.4	1853	1	MY5A_MOUSE	O09110	mus musculu
309	24	77.4	637	1	PRIM_LACLA	O04505	lactococcus	382	24	77.4	1855	1	MY5A_HUMAN	O09110	homo sapien
310	24	77.4	638	1	YOGS_BACST	P54496	bacillus su	383	24	77.4	1872	1	T2D1_MOUSE	P21678	homo sapien
311	24	77.4	664	1	UYRB_THEMA	O92482	thermotoga	384	24	77.4	2206	1	POLG_CXA21	P22055	c genome po
312	24	77.4	668	1	DNJL_BACST	O34398	bacillus su	385	24	77.4	2291	1	RRPB_BEV	P22055	c genome po
313	24	77.4	670	1	CUI3_HUMAN	O95447	homo sapien	386	24	77.4	2704	1	BPAL_HUMAN	O03001	homo sapien
314	24	77.4	685	1	RNT1_DROME	O9uy43	drosophila	387	24	77.4	2748	1	NUM1_YEAST	O00402	saccharomyc
315	24	77.4	686	1	RECG_BORBU	O85d28	borrelia bu	388	24	77.4	2871	1	FBNI_BOVIN	P08133	bos taurus
316	24	77.4	693	1	RECG_ECOS7	O85d68	escherichia	389	24	77.4	2871	1	FBNI_HUMAN	P35555	homo sapien
317	24	77.4	693	1	RECG_ECOLI	P24230	escherichia	390	24	77.4	2871	1	FBNI_PIG	Q9tvt8	sus scrofa
318	24	77.4	701	1	HS90_PODAN	O43109	podospira a	391	24	77.4	3210	1	CENF_HUMAN	P27395	homo sapien
319	24	77.4	709	1	CDAT_PLAFA	P46468	plasmodium	392	24	77.4	3432	1	POLG_JAEVI	P27395	genome po
320	24	77.4	713	1	HS9B_HORSE	O094x8	equus cabal	393	24	77.4	3432	1	POLG_JAEV5	P19110	genome po
321	24	77.4	716	1	HS83_DROAV	O02192	drosophila	394	24	77.4	3587	1	SRR2_JAEVJ	P03286	genome po
322	24	77.4	717	1	HS83_DROME	O08828	drosophila	395	24	77.4	3587	1	FIXS_RHIME	O04747	bacillus su
323	24	77.4	719	1	HS9A_HORSE	O96kx7	equus cabal	396	24	77.4	61	1	Y063_ARCTU	O30177	archaeoglob
324	24	77.4	723	1	HS9B_HUMAN	P08238	homo sapien	397	24	77.4	61	1	Y063_ARCTU	O30177	archaeoglob
325	24	77.4	723	1	HS9B_MOUSE	P11499	mus musculu	398	24	77.4	63	1	Y063_ARCTU	P08366	escherichia

399	23	74.2	67	1	RFL_STRGC	P47850 streptococc	472	23	74.2	284	1	SA41_MOUSE	Q9d597 mus musculu
400	23	74.2	85	1	VG73_BPM15	O05288 mycobacteri	473	23	74.2	284	1	TPM1_SCHMA	P42637 schistosoma
401	23	74.2	87	1	R15_TOBAC	P63133 nicotiana t	474	23	74.2	290	1	ALF2_BACSU	P42420 bacillus su
402	23	74.2	88	1	R15_ARATH	P56805 arabidopsis	475	23	74.2	295	1	OTC_HALNI	O48290 halobacteri
403	23	74.2	88	1	V176_TREPA	O83206 treponema p	476	23	74.2	305	1	CMW2_SCHPO	O94417 schizosacch
404	23	74.2	90	1	R15_SPTOL	O09m34 spiniacia ol	477	23	74.2	305	1	YES5_METUA	O58650 methanococc
405	23	74.2	96	1	NEUT_BRARE	Q31893 brachydanio	478	23	74.2	308	1	AMID_STEPPN	P18194 streptococc
406	23	74.2	96	1	NEUT_CARAU	P28672 carassius a	479	23	74.2	309	1	WR26_ARATH	P51593 rattus norv
407	23	74.2	102	1	VNI6_YEAST	P28262 saccharomyc	480	23	74.2	310	1	URB1_RAT	P52336 flaveria bi
408	23	74.2	109	1	Y819_METUA	O58229 methanococc	481	23	74.2	311	1	F35F_FLACH	P52336 flaveria bi
409	23	74.2	120	1	R117_BACHD	O50655 bacillus ha	482	23	74.2	312	1	F35F_FLACH	P52336 flaveria bi
410	23	74.2	121	1	RK14_SPTOL	P09556 spiniacia ol	483	23	74.2	314	1	YFCH_HAEIN	P45106 haemophilus
411	23	74.2	122	1	RK14_ARATH	P56792 arabidopsis ol	484	23	74.2	315	1	P2A_PARTE	P48726 parametium
412	23	74.2	123	1	RK14_MATZE	P08559 zea mays (m	485	23	74.2	317	1	OMP7_ECO57	P58603 escherichia
413	23	74.2	123	1	RK14_ORYA	P12137 oryza sativ	486	23	74.2	317	1	OMP7_ECO57	P58603 escherichia
414	23	74.2	123	1	RK14_TOBAC	P63342 nicotiana t	487	23	74.2	321	1	DE76_CHLPR	O9x1y6 chloroella p
415	23	74.2	124	1	CD59_RABIT	O77541 cycloclagus	488	23	74.2	322	1	XCX1_MOUSE	O9x1m1 mus musculu
416	23	74.2	126	1	YD43_MYCLE	P54144 mycobacteri	489	23	74.2	322	1	YOUN_BACSU	P54338 bacillus su
417	23	74.2	126	1	YD43_MYCGR	O11013 mycobacteri	490	23	74.2	329	1	I329_ASEB7	P27455 african swi
418	23	74.2	136	1	R117_RICPR	O9x2c0 rickettsia	491	23	74.2	334	1	OOR_YEAST	P38230 saccharomyc
419	23	74.2	137	1	PLAV_MEGFL	P00331 megasphaera	492	23	74.2	336	1	RPOA_AGR15	O9710 agrobacteri
420	23	74.2	141	1	RFL_ECOLI	P28359 escherichia	493	23	74.2	336	1	RPOA_RH10	O98133 rhizobium l
421	23	74.2	157	1	YB23_PYRAB	Q9ux24 pyrococcus	494	23	74.2	336	1	RPOA_RH10	O92522 rhizobium m
422	23	74.2	161	1	YY22_MYCLE	Q49864 mycobacteri	495	23	74.2	337	1	RPOA_RH10	O92522 rhizobium m
423	23	74.2	161	1	IPYR_BACP3	P19514 bacillus ps	496	23	74.2	337	1	RPOA_BRUME	O916c7 bartonella
424	23	74.2	164	1	IPYR_BACST	O05724 bacillus st	497	23	74.2	338	1	Y941_METUA	O57711 methanococc
425	23	74.2	166	1	VE07_VACCV	P21048 vaccinia st	498	23	74.2	340	1	RPOA_RICCN	O92420 rickettsia
426	23	74.2	176	1	IPYR_GLUOX	O05545 gluconobact	499	23	74.2	340	1	RPOA_RICPR	O92659 rickettsia
427	23	74.2	176	1	RIMW_THEMA	Q23164 thermotoga	500	23	74.2	341	1	ARGC_METUA	O92659 rickettsia
428	23	74.2	185	1	PAGC_SALTU	P33988 salmonella	501	23	74.2	341	1	MAG_DERFA	O58496 methanococc
429	23	74.2	185	1	RPO6_MXYVL	P18620 myxoma viru	502	23	74.2	347	1	DHAS_LEGPN	P39673 dematophag
430	23	74.2	186	1	RPO6_FOWPV	Q85280 fowlpox vir	503	23	74.2	347	1	H806_METUA	O31219 leijonella
431	23	74.2	188	1	Y4AO_RHISM	P55364 rhizobium s	504	23	74.2	348	1	AROH_ERWHE	O54459 erwinia her
432	23	74.2	195	1	NUGM_ACACA	O37383 acanthamoeb	505	23	74.2	348	1	VMAT_SENDS	P27667 sendai viru
433	23	74.2	198	1	PBD2_CHICK	O73888 gallus gall	506	23	74.2	348	1	VMAT_SENDS	P17748 sendai viru
434	23	74.2	198	1	PBD2_CHICK	Q5j3h7 mus musculu	507	23	74.2	348	1	VMAT_SENDS	P03326 sendai viru
435	23	74.2	198	1	PBD2_MOUSE	O55343 rattus norv	508	23	74.2	348	1	VMAT_SENDS	P06446 sendai viru
436	23	74.2	201	1	MUSC_MOUSE	O88940 mus musculu	509	23	74.2	349	1	AOX1_SAUGU	P12185 saurotatum
437	23	74.2	204	1	VATD_BORBU	O51119 borrelia bu	510	23	74.2	351	1	CBF1_YEAST	P17106 saccharomyc
438	23	74.2	207	1	RUD2_BUCAL	P57053 buchiera ap	511	23	74.2	351	1	VORH_METTH	O26800 methanobact
439	23	74.2	213	1	YIN2_YEAST	P40465 saccharomyc	512	23	74.2	353	1	HME1_DROME	P09089 drosophila
440	23	74.2	215	1	TRJ9_ECOLI	Q00738 escherichia	513	23	74.2	353	1	ILB4_RABIT	P21109 oryctolagus
441	23	74.2	216	1	FUCA_HAEIN	P44777 haemophilus	514	23	74.2	357	1	RFL_LACLA	O9c1x3 lactococcus
442	23	74.2	218	1	R110_DROME	O61231 drosophila	515	23	74.2	358	1	PHLC_TREBB	O9c1x3 trypanosoma
443	23	74.2	221	1	SFR9_HUMAN	Q13242 homo sapien	516	23	74.2	359	1	RFL_STRAG	O9a229 streptococc
444	23	74.2	224	1	KCY_THETN	O8ra78 thermoaenar	517	23	74.2	359	1	RFL_STRPN	O97120 streptococc
445	23	74.2	224	1	E681_CHLPH	O927m3 chlamydia p	518	23	74.2	359	1	RFL_STRPY	O99425 streptococc
446	23	74.2	232	1	E4RE_METH	O26350 methanobact	519	23	74.2	360	1	ILB8_BOVIN	O28603 bos taurus
447	23	74.2	232	1	YB9H_YEAST	P38340 saccharomyc	520	23	74.2	360	1	RFL_PSEAE	P42806 pseudomonas
448	23	74.2	237	1	VD03_VACCC	P21009 vaccinia vi	521	23	74.2	361	1	MK14_CYPCA	O90336 cyprinus ca
449	23	74.2	237	1	VD03_VACCV	P43302 vaccinia vi	522	23	74.2	361	1	RFL_HAEIN	P43917 haemophilus
450	23	74.2	237	1	VD03_VARY	P33068 variola vir	523	23	74.2	366	1	H182_HAEIN	O57004 haemophilus
451	23	74.2	238	1	YFBN_ECOLI	P76484 escherichia	524	23	74.2	368	1	DXR_HELPJ	O9m16 helicobacte
452	23	74.2	243	1	GNTB_BACSU	P10585 bacillus su	525	23	74.2	368	1	SNCL_HUMAN	O51533 homo sapien
453	23	74.2	244	1	WC23_VACCC	P21080 vaccinia vi	526	23	74.2	370	1	K2C8_BOVIN	P05786 bos taurus
454	23	74.2	250	1	TDXH_AERPE	O9y910 aeropyrum p	527	23	74.2	372	1	CARA_CAMDE	O9pmg8 campylobact
455	23	74.2	253	1	HIS6_PYRAE	P82y16 pyrobaculum	528	23	74.2	373	1	HME1_DROSU	O24648 drosophila
456	23	74.2	253	1	VC23_VARY	O8y016 variola vir	529	23	74.2	374	1	YG09_METUA	O59004 methanococc
457	23	74.2	257	1	ICOE_MOUSE	O89094 mus musculu	530	23	74.2	375	1	YHR4_PSESY	P20400 pseudomonas
458	23	74.2	258	1	VC23_VACCC	P19063 vaccinia vi	531	23	74.2	384	1	O46B_DROME	O9y3n2 drosophila
459	23	74.2	261	1	TNE5_BOVIN	P51749 bos taurus	532	23	74.2	389	1	AAT_SYNY3	O55128 synechocyst
460	23	74.2	262	1	TRPA_AQUAE	O67502 aquifex aeo	533	23	74.2	393	1	PGIR_SYNY3	P48979 pinus pels
461	23	74.2	264	1	TYSY_RALSO	O8y0u6 ralsomonia s	534	23	74.2	394	1	EFPU_BORBU	P50062 borrelia bu
462	23	74.2	265	1	HME2_BRARE	P09015 brachydanio	535	23	74.2	398	1	PSMR_ARCFU	O28303 archaoglob
463	23	74.2	270	1	YNE8_YEAST	P53722 saccharomyc	536	23	74.2	398	1	Y4QJ_RHISM	P55631 rhizobium s
464	23	74.2	271	1	PUNA_BACSU	P46354 bacillus su	537	23	74.2	400	1	YD01_METUA	REP4_AGRRH
465	23	74.2	272	1	VG15_BP14	P11112 bacterioph	538	23	74.2	404	1	REP4_AGRRH	O56692 methanococc
466	23	74.2	276	1	PRXC_SYNY3	O55921 synechocyst	539	23	74.2	406	1	OLDA_DROME	P05682 agrobacteri
467	23	74.2	278	1	UL07_EBY	P29882 epstein-bar	540	23	74.2	410	1	FBW4_MOUSE	O9y221 drosophila
468	23	74.2	279	1	YEB8_YEAST	P39995 saccharomyc	541	23	74.2	411	1	INTR_ECOLI	O9ymj2 mus musculu
469	23	74.2	282	1	TR2A_HUMAN	O13595 homo sapien	542	23	74.2	412	1	FBW4_HUMAN	P76056 escherichia
470	23	74.2	283	1	RUPAP_SOYBN	P39657 glycine max	543	23	74.2	413	1	DCHS_LYCES	P54772 lyopersico
471	23	74.2	284	1	SA41_HUMAN	O9br01 homo sapien	544	23	74.2	413	1	DP41_BACHD	O9kc07 bacillus ha

837	22	71.0	162	1	IPYR_MYCTU	006379 mycobacteri	910	22	71.0	227	1	COX2_BUBDE	P50678 bubalus dep
838	22	71.0	167	1	NDOE_RICCN	Q931d9 rickettsia	911	22	71.0	227	1	COX2_CANAD	O47667 canis adust
839	22	71.0	168	1	GA15_CRILC	P14607 cricetulus	912	22	71.0	227	1	COX2_CANAD	O47668 canis aureu
840	22	71.0	168	1	GA15_MOUSE	P35639 mus musculu	913	22	71.0	227	1	COX2_CANFA	O63855 canis famli
841	22	71.0	169	1	GA15_HUMAN	P35638 homo sapien	914	22	71.0	227	1	COX2_CANME	O47671 canis mesom
842	22	71.0	173	1	CD3D_RAT	P19377 rattus norv	915	22	71.0	227	1	COX2_CANST	P80831 canis simen
843	22	71.0	174	1	CU22_HOMMO	O02388 bombyx mori	916	22	71.0	227	1	COX2_CAPII	O37430 capra hircu
844	22	71.0	174	1	GRPE_METHH	O02350 methanobact	917	22	71.0	227	1	COX2_CERST	O03851 ceratotheri
845	22	71.0	175	1	CYT2_STOHE	P07845 stoichiactis	918	22	71.0	227	1	COX2_CERNU	O37437 cervus unic
846	22	71.0	175	1	NIDM_BOVIN	O02373 bos taurus	919	22	71.0	227	1	COX2_CHRRR	O47670 chrysocyon
847	22	71.0	176	1	MOC_HELPU	O92145 helicobacte	920	22	71.0	227	1	COX2_CRABU	P50685 cratogeomys
848	22	71.0	176	1	MOC_HELPU	P56421 helicobacte	921	22	71.0	227	1	COX2_CRACA	P50686 cratogeomys
849	22	71.0	177	1	FABA_PASMU	Q9cne8 pasteurella	922	22	71.0	227	1	COX2_CUCAL	O47668 cuon alpinu
850	22	71.0	177	1	NDOE_RICPR	O9cdh5 rickettsia	923	22	71.0	227	1	COX2_DAMP	P50679 damaliscus
851	22	71.0	177	1	VC08_YACCV	P17364 vaccinia vi	924	22	71.0	227	1	COX2_DASNO	P50687 dasypus nov
852	22	71.0	178	1	YEDR_ECOCI	P76514 escherichia	925	22	71.0	227	1	COX2_DUSRH	O47672 duslicyon th
853	22	71.0	179	1	UECB_XENLA	P56516 xenopus lae	926	22	71.0	227	1	COX2_EQUUS	P92478 equus asinu
854	22	71.0	179	1	UBCC_HUMAN	O00762 homo sapien	927	22	71.0	227	1	COX2_FELCA	P48890 felis silve
855	22	71.0	181	1	TCIP_BRUMA	P09697 brugia mala	928	22	71.0	227	1	COX2_GAZSP	P48890 gazella spe
856	22	71.0	181	1	TCIP_MOCBA	Q962a2 wuchereria	929	22	71.0	227	1	COX2_HALGR	P38596 halichoerus
857	22	71.0	181	1	VHL_MOUSE	P40338 mus musculu	930	22	71.0	227	1	COX2_HIPAM	Q92zy9 hippopotamu
858	22	71.0	184	1	VC08_YACCC	P21041 vaccinia vi	931	22	71.0	227	1	COX2_HORSE	P48660 equus cabal
859	22	71.0	184	1	YARO_ACTPL	P46393 actinobacil	932	22	71.0	227	1	COX2_LEMCA	P47674 lemur carita
860	22	71.0	188	1	AMPN_METFE	P22624 methanother	933	22	71.0	227	1	COX2_LYGPI	O47674 lycan pict
861	22	71.0	195	1	RM1_CAREL	P34264 caenorhabdi	934	22	71.0	227	1	COX2_MACRO	P32662 macrodon pict
862	22	71.0	198	1	MOBA_CLOPE	Q9wx94 clostridium	935	22	71.0	227	1	COX2_MITCE	P24988 microtus pe
863	22	71.0	198	1	PCB2_HUMAN	O60760 homo sapien	936	22	71.0	227	1	COX2_MOUSE	P04005 mus musculu
864	22	71.0	202	1	HPRT_MYCTU	O06383 mycobacteri	937	22	71.0	227	1	COX2_NYCP	P38596 nyctereutes
865	22	71.0	203	1	HPRT_MYCAV	P96794 mycobacteri	938	22	71.0	227	1	COX2_PHOVT	O00528 phoca vitul
866	22	71.0	203	1	HPRT_MYCLE	O65337 mycobacteri	939	22	71.0	227	1	COX2_PROPA	P86042 propithecus
867	22	71.0	206	1	AFR_PIG	P82125 sus scrofa	940	22	71.0	227	1	COX2_PSEBX	O47674 pseudalopex
868	22	71.0	206	1	UPP_MYCGE	P47276 mycoplasma	941	22	71.0	227	1	COX2_PSEBX	O47674 pseudalopex
869	22	71.0	207	1	CB23_TETTH	P20473 tetrahymena	942	22	71.0	227	1	COX2_PSEBU	O47674 pseudalopex
870	22	71.0	208	1	ETFS_CYACA	Q9um32 cyamidum c	943	22	71.0	227	1	COX2_RABIT	P80404 ryctolagus
871	22	71.0	208	1	Y24K_MOUSE	O09388 mus musculu	944	22	71.0	227	1	COX2_RAT	O60408 rattus norv
872	22	71.0	209	1	VAID_METMA	O60188 methanosarc	945	22	71.0	227	1	COX2_RH1UN	P06190 rhinoceros
873	22	71.0	209	1	YJ9B_YEAST	P47165 saccharomyc	946	22	71.0	227	1	COX2_SICCA	O50691 sciurus car
874	22	71.0	211	1	GT28_SCHBO	P30113 schistosoma	947	22	71.0	227	1	COX2_SHEEP	O47674 speothos ve
875	22	71.0	211	1	GT28_SCHBO	P30114 schistosoma	948	22	71.0	227	1	COX2_SHEEP	P50675 synderus ca
876	22	71.0	211	1	VHL_HUMAN	P40337 homo sapien	949	22	71.0	227	1	COX2_SINCA	P80404 tarsius ban
877	22	71.0	215	1	I12A_MOUSE	P44331 mus musculu	950	22	71.0	227	1	COX2_TARBA	P80404 tarsius syr
878	22	71.0	216	1	COX2_CALGO	P30661 callinico g	951	22	71.0	227	1	COX2_TARSY	O37685 tragelaphus
879	22	71.0	217	1	VNS3_AHSV3	P33864 african hor	952	22	71.0	227	1	COX2_TRAIM	O47680 vulpes macr
880	22	71.0	218	1	CB25_TETTH	P09226 tetrahymena	953	22	71.0	227	1	COX2_VU1MA	O47680 vulpes macr
881	22	71.0	219	1	I12A_CERTO	P46661 cercocebus	954	22	71.0	227	1	COX2_VU1VU	O47681 vulpes vulp
882	22	71.0	219	1	I12A_HUMAN	P29459 homo sapien	955	22	71.0	227	1	COX2_VU1ZE	O47673 vulpes zerd
883	22	71.0	219	1	I12A_MACMU	P48091 macaca mula	956	22	71.0	228	1	COX2_MOUSE	P55330 mus musculu
884	22	71.0	219	1	Y111_STRCO	O86576 streptomyc	957	22	71.0	228	1	COX2_PIG	P30667 sus scrofa
885	22	71.0	220	1	NDKM_DICDI	P36093 dictyosteli	958	22	71.0	228	1	RPE_STRCO	O31025 streptomyc
886	22	71.0	220	1	YTFE_ECOCI	P33313 escherichia	959	22	71.0	229	1	C79B_HUMAN	O40259 homo sapien
887	22	71.0	220	1	YFBB_BACSV	P37502 bacillus su	960	22	71.0	229	1	COX2_XENLA	P00407 xenopus lae
888	22	71.0	221	1	I12A_BOVIN	P54349 bos taurus	961	22	71.0	230	1	COX2_GADMO	O37741 gadus morhu
889	22	71.0	221	1	I12A_CERET	O28233 cervus elap	962	22	71.0	230	1	COX2_ONCMY	P48171 oncorhynch
890	22	71.0	222	1	I12A_CANFA	O28267 canis famli	963	22	71.0	230	1	COX2_ORNAN	O37718 ornithorhyn
891	22	71.0	222	1	I12A_FELCA	O02743 felis silve	964	22	71.0	230	1	COX2_SALSA	Q37677 salmo salar
892	22	71.0	222	1	I12A_HORSE	O9xsq6 equus cabal	965	22	71.0	230	1	COX2_SCYCA	O79404 scylliorhnu
893	22	71.0	222	1	I12A_PIG	O29053 sus scrofa	966	22	71.0	230	1	COX2_SQUAC	O32551 squatus aca
894	22	71.0	222	1	OS28_PLAGA	O05439 plasmodium	967	22	71.0	230	1	EBP_HUMAN	O15125 homo sapien
895	22	71.0	222	1	Y434_METLA	O57876 methanococ	968	22	71.0	231	1	COX2_ALOPA	P81513 alouatta pa
896	22	71.0	223	1	I12A_MARMO	O61728 marmota mon	969	22	71.0	234	1	PUR7_PYRAE	Q628K3 pyrodictum
897	22	71.0	223	1	YKPL_KLUTLA	P05467 kluyveromyc	970	22	71.0	234	1	Y491_ARCFU	O39759 arctaeoglob
898	22	71.0	225	1	URK_HALNI	Q9hgq9 halobacteri	971	22	71.0	235	1	1433_XENLA	P29309 xenopus lae
899	22	71.0	225	1	YFGL_YEAST	P43543 saccharomyc	972	22	71.0	237	1	SHR5_YEAST	P11912 saccharomyc
900	22	71.0	226	1	COX2_PERFA	O37595 perognathus	973	22	71.0	238	1	AGRA_STRAN	P13131 staphylococ
901	22	71.0	227	1	COX2_ACOMT	P50672 acornys wills	974	22	71.0	238	1	UBIG_NEIMA	O37467 neisseria m
902	22	71.0	227	1	COX2_ANTAM	Q37369 antilocapra	975	22	71.0	238	1	COX2_PSEBX	O37467 neisseria m
903	22	71.0	227	1	COX2_APOSY	P50673 apodemus sy	976	22	71.0	239	1	HIS4_MERTH	O68931 metanodact
904	22	71.0	227	1	COX2_SALMU	P41294 balaenopter	977	22	71.0	240	1	NOO2_PARDE	P29914 paracoccus
905	22	71.0	227	1	COX2_SALPH	P24986 balaenopter	978	22	71.0	240	1	LEF5_GVCL	P39717 cryptophleb
906	22	71.0	227	1	COX2_BISBI	Q37416 bison bison	979	22	71.0	240	1	PYRH_BACSV	O31749 bacillus su
907	22	71.0	227	1	COX2_BOSJA	P50674 bos javanic	980	22	71.0	241	1	Y136_ARATH	O31749 arabidopsis
908	22	71.0	227	1	COX2_BOSMR	O37419 boselaphus	981	22	71.0	241	1	FLGF_RHILLO	O98ba3 rhizobium l
909	22	71.0	227	1	COX2_BOVIN	P00404 bos taurus	982	22	71.0	241	1	GRS_MUSDO	P46437 musca domes

ALIGNMENTS

PSMR_METUA	STANDARD	PRT	430 AA
983	22	71.0	241
984	22	71.0	244
985	22	71.0	245
986	22	71.0	245
987	22	71.0	245
988	22	71.0	245
989	22	71.0	245
990	22	71.0	245
991	22	71.0	245
992	22	71.0	245
993	22	71.0	245
994	22	71.0	245
995	22	71.0	245
996	22	71.0	245
997	22	71.0	245
998	22	71.0	245
999	22	71.0	246
1000	22	71.0	246

ALIGNMENTS

RESULT 1

PSMR_METUA STANDARD; PRT; 430 AA.

AC 058576; (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteasome-activating nucleotidase (Proteasome regulatory subunit).
 GN PAN OR M1176.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utermack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Borodovsky M.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 [2]
 RP CHARACTERIZATION.
 RX MEDLINE=20158885; PubMed=10692374;
 RA Wilson H.L., Ou M.S., Aldrich H.C., Maupin-Furlow J.,
 RT "Biochemical and physical properties of the Methanococcus jannaschii
 20S proteasome and PAN, a homolog of the Atpase (Rpt) subunits of the
 eucaryal 26S proteasome.";
 RL J. Bacteriol. 182:1680-1692(2000).
 [3]
 RP CHARACTERIZATION.
 RX MEDLINE=99403035; PubMed=10473546;
 RA Zwickl P., Ng D., Woo K.M., Klenk H.-P., Goldberg A.L.,
 RT "An archaeobacterial Atpase, homologous to Atpases in the eukaryotic 26
 S proteasome, activates protein breakdown by 20 S proteasomes.";
 RL J. Biol. Chem. 274:26008-26014(1999).
 CC -1- FUNCTION: Required for the ATP- or CTP-dependent degradation of
 CC proteins, but not small peptides, by the 20S proteasome.
 CC -1- COFACTOR: MAGNESIUM.
 CC -1- ENZYME REGULATION: INHIBITED BY EDTA, NEM AND POMS.
 CC -1- SUBUNIT: HOMOHexamer (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: ITS OPTIMUM PH IS 7-8 AND THE OPTIMUM TEMPERATURE

IS 80 DEGREES CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U67559; AAB99179.1; -
 DR TIGR: M1176; -
 DR InterPro: IPR003593; AAA_Atpase.
 DR InterPro: IPR003959; AAA_Atpase_centrl.
 DR InterPro: IPR003960; AAA_sub.
 DR Pfam: PF00004; AAA.1.
 DR SMART: SM00382; AAA.1.
 DR PROSITE: PS00674; AAA.1.
 DR TIGRfams: TIGR01242; 26Sp45.1.
 KW Proteasome; ATP-binding; Complete proteome.
 FT NE_BIND 211 218 ATP (POTENTIAL).
 SQ SEQUENCE 430 AA; 48689 MW; 3FD2E94A68D463DD CRC64;

Query Match 93.5%; Score 29; DB 1; Length 430;
 Best Local Similarity 83.3%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDD 6
 DB 167 VREDD 172

RESULT 2

WIS4_SCHPO STANDARD; PRT; 1401 AA.

AC 014299; P87062; Q92384;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MAP kinase kinase kinase wis4 (EC 2.7.1.1.) (MAP kinase kinase kinase
 DE waki) (MAP kinase kinase kinase waki).
 GN WIS4 OR WAKI OR WIKI OR SPAC9G1.02.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98026892; PubMed=9321395;
 RA Samejima I., Mackie S., Fantes P.A.,
 RT "Multiple modes of activation of the stress-responsive MAP kinase
 RT pathway in fission yeast.";
 RL EMBO J. 16:6162-6170(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottier S.,
 RA Calbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Kochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*,"
 RL Nature 415:871-880(2002).
 RN [3]
 RP SEQUENCE OF 96-1401 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97282620; PubMed=9136929;
 RA Shieh J.C., Wilkinson M., Buck V., Morgan B., Makino K.,
 RA Miller J.B.A.;
 RT "The Mcs4 response regulator coordinately controls the stress-
 RT activated Wsk1-Wsk1-Styl1 MAP kinase pathway and fission yeast cell
 RT cycle.";
 RL Genes Dev. 11:1008-1022(1997).
 RN [4]
 RP SEQUENCE OF 457-543 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97331762; PubMed=9188094;
 RA Shiozaki K., Shiozaki M., Russell P.,
 RT "Mcs4 mitotic catastrophe suppressor regulates the fission yeast cell
 RT cycle through the Wsk1-Wsk1-Styl1 kinase cascade.";
 RL Mol. Biol. Cell 8:409-419(1997).
 CC -1- FUNCTION: INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS
 CC ACTIVATED IN UNDER CONDITIONS OF HEAT SHOCK, OXIDATIVE STRESS OR
 CC LIMITED NUTRITION. UNLIKE WSK1, IT IS NOT ACTIVATED BY CHANGES IN
 CC THE OSMOLARITY OF THE EXTRACELLULAR ENVIRONMENT. ACTIVATES THE
 CC WSK1 MAP KINASE KINASE BY PHOSPHORYLATION.
 CC SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: 298763; CA011500.1; -
 DR EMBL: Y07760; CA069030.1; ALT-INT.
 DR EMBL: Y11969; CA02718.1; -
 DR EMBL: U81521; AAB39762.1; -
 DR HSSP: P24941; ICRP.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 KW Transferrase: Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 1037 1306 PROTEIN KINASE.
 FT NP_BIND 1043 1051 ATP (BY SIMILARITY).
 FT BINDING 1066 1066 ATP (BY SIMILARITY).
 FT ACT_SITE 1161 1161 BY SIMILARITY.
 FT CONFLICT 484 485 RL -> SP (IN REF. 3).
 FT CONFLICT 1346 1346 D -> V (IN REF. 1).
 SQ SEQUENCE 1401 AA; 160537 MW; 529A989D2B627F3D CRC64;

DB 1194 VREDI 1199
 |||||:
 DB 1194 VREDI 1199
 RESULT 3
 ID RPOC_TREPA STANDARD; PRI: 1416 AA.
 AC 083270;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
 DE beta' chain) (RNA polymerase beta' subunit).
 GN RPOC OR TP0242.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_Taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=96832770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback J.,
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis
 RT spirochete.";
 RL Science 281:375-388(1998).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC [RNA(N)].
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: AE001205; AAC65230.1; -
 DR HSSP: O9KW06; IHQM.
 DR TIGR: TP0242.
 DR InterPro: IPR000722; RNA_POL_A.
 DR InterPro: IPR002879; RNA_POL_A2.
 DR Pfam: PF00623; RNA_POL_A.1.
 DR Pfam: PF01854; RNA_POL_A2.2.
 KW Transferrase: DNA-directed RNA polymerase; Transcription;
 KW complete proteome.
 SQ SEQUENCE 1416 AA; 159789 MW; 780DEB6A1CAF7C9C CRC64;

Query Match 93.5%; Score 29; DB 1; Length 1401;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDI 6

DB 1052 VREDI 1057

RESULT 4
 ID CHIS_BACSU STANDARD; PRI: 277 AA.
 AC 007921;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chitosanase precursor (EC 3.2.1.132).
 CSN.
 OS Bacteria subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI-TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97286548; PubMed=9141695;
 RA Paro V., San Roman M., Gallardo I., Purnelle B., Bolotin A.,
 RA Sorokin A., Mellado R.P.: The Bacillus subtilis genome comprising genes
 RA "A 2391 bp region of the Bacillus subtilis genome comprising genes
 RA located upstream and downstream of the lev operon."
 RL Microbiology 143:1321-1326(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97453479; PubMed=9308178;
 RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
 RA Duesthoeft A., Ehrlich S.D.:
 RA "Sequence of the Bacillus subtilis genome region in the vicinity of
 RA the lev operon reveals two new extracytoplasmic function RNA
 RA polymerase sigma factors SigV and SigZ."
 RT Microbiology 143:2939-2943(1997).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berrero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denlioz F., Devine K.M., Duesthoeft A., Ehrlich S.D., Emmerson P.T.,
 RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haeche J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya N., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klierer-Bianchard M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portelle D., Portwollik S., Prescott S.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roeha E., Roche B., Rose M., Sadele Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska E., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni K.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viati A., Wandut R., Wedler E., Wedler H., Weitenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis."
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: AID IN THE DEFENSE TO INVADING FUNGAL PATHOGENS BY
 CC DEGRADING THEIR CELL WALL CHITOSAN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of beta-1,4-linkages between N-
 CC acetyl-D-glucosamine and D-glucosamine residues in a partly
 CC acetylated chitosan.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 46 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X92868; CAA63455.1;
 CC DR EMBL: U93875; AAB80882.1;
 CC DR EMBL: 299117; CAB14630.1;
 CC DR HSSP: P33665; ICHK.
 CC DR SUBLIST: BG11923; csn.
 CC DR InterPro: IPR00400; GH_46.
 CC DR Pfam: PF01374; Glyco_hydro_46; 1.
 CC DR PROSITE: PS60000; CHITOSANASE_46_80; 1.
 CC KM Hydrolyase; Glycosidase; Signal; Complete proteome.
 CC FT SIGNAL 1 35
 CC FT CHAIN 36 277
 CC FT ACT_SITE 54 54
 CC FT ACT_SITE 70 70
 CC FT ACT_SITE 70 70
 CC SQ SEQUENCE 277 AA: 31496 MW: 8C17C156AE27F781 CRC64;
 QY 1 VRYEDL 6
 Db 222 VRYDDL 227
 Query Match 90.3%; Score 28; DB 1; Length 277;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 RESULT 5
 CAC3_STRPN STANDARD: PRT: 306 AA.
 ID CAC3_STRPN
 AC 054800; Q54612;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) (UDP-glucose
 DE pyrophosphorylase) (UDPgc) (Alpha-D-glucosyl-1-phosphate
 DE uridylyltransferase) (uridine diphosphoglucose pyrophosphorylase).
 DE CAP3 OR CP83U.
 GN Streptococcus pneumoniae.
 OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI-TaxID=1313;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=406 / Type 3;
 RX MEDLINE=96144241; PubMed=8566758;
 RA Arcubielta C., Garcia E., Lopez R.;
 RT "Sequence and transcriptional analysis of a DNA region involved in
 RT the production of capsular polysaccharide in Streptococcus pneumoniae
 RT type 3."
 RL Gene 167:1-7(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WD2 / Serotype 3;
 RX MEDLINE=95173611; PubMed=7869055;
 RA Dillard J.P., Vandersa M.M., Vother J.;
 RT "Characterization of the cassette containing genes for type 3
 RT capsular polysaccharide biosynthesis in Streptococcus pneumoniae."
 RL J. Exp. Med. 181:973-983(1995).
 CC -1- CATALYTIC ACTIVITY: UTP + alpha-D-glucose 1-phosphate =
 CC diphosphate + UDP-glucose.
 CC -1- PATHWAY: PRODUCTION OF CAPSULAR POLYSACCHARIDE.
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC UDPgc FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: Z47210; CAAG7405.1; -
DR EMBL: U15171; AAC43313.1; -
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00483; NTP_transferase; 1.
DR TIGRfams: TIGR01099; galU; 1.
KW TRANSFERASE; Kinase; Nucleotidyltransferase.
FT CONFLICT 204 204 N -> H (IN REF. 2).
SQ SEQUENCE 306 AA; 33919 MW; F8E243D00F47EEA8 CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DA 1; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 165 VRYEDV 170

RESULT 6
ASRC_SALTY STANDARD; PRT; 337 AA.
ID ASRC_SALTY STANDARD; PRT; 337 AA.
AC P26476;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anaerobic sulfite reductase subunit C (EC 1.8.1.-).
GN ASRC OR STM2550 OR STY2796.
OS Salmonella typhimurium, and
OC Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhimurium; STRAIN-EB303;
RA MEDLINE=91139599; PubMed=1704886;
RA Huang C.J., Barrett E.L.;
RT *Sequence analysis and expression of the Salmonella typhimurium asr
RT operon encoding production of hydrogen sulfide from sulfite.*;
RL J. Bacteriol. 173:1544-1553(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhimurium; STRAIN-LT2 / SGGSC1412 / ATCC 700720;
RA MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT *Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.*;
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhi; STRAIN-CT18;
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Doud T., White N., Farrar J.,
RA Felwell I., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT *Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.*;
RL Nature 413:848-852(2001).
RN [4]
RP FUNCTION: THIS ENZYME CATALYZES THE HYDROGEN SULFIDE PRODUCTION
RC FROM SULFITE. IT IS STRICTLY ANAEROBIC. IT IS REGULATED BY
CC ELECTRON ACCEPTORS RATHER THAN BY CYSTEINE.
CC -1- CATALYTIC ACTIVITY: Hydrogen sulfide + 3 NAD(+) + 3 H(2)O =
CC sulfite + 3 NADH.
CC NCBI_TaxID=1773;

```

```

CC -----
CC COFACTOR: THIS PROTEIN CONTAINS TWO SIROHEME AND EIGHT 4FE-4S
CC IRON-SULFUR CENTER AS PROSTHETIC GROUPS.
CC -1- PATHWAY: Sulfite reduction.
CC -1- SUBUNIT: THE ANAEROBIC SULFITE REDUCTASE SEEMS TO CONSIST OF THREE
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- INDUCTION: BY SULFITE.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
CC FOUND IN NITRATE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
CC SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57706; AAA99277.1; -
DR EMBL: AE008816; AAL21444.1; -
DR EMBL: AL627275; CAD02753.1; -
DR PIR: C38453; C38453.
DR StyGene; SG10023; asrc.
DR InterPro: IPR001450; 4fe4s_ferredoxin.
DR InterPro: IPR000660; Ntr_sir.
DR InterPro: IPR005117; Ntr_sir_fer.
DR Pfam: PF00037; fer4_2.
DR Pfam: PF01077; Ntr_sir; 1.
DR Pfam: PF03460; Ntr_sir_fer; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
DR PROSITE; PS00365; Ntr_sir; 1.
KW Oxidoreductase; Electron transport; Heme; NAD; Iron-sulfur; 4fe-4s;
KW Complete proteome.
FT METAL 115 115 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 121 121 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 153 153 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 157 157 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 180 180 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 183 183 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 186 186 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 190 190 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 212 212 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 215 215 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 218 218 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 222 222 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT CONFLICT 247 250 TPRV -> SGAL (IN REF. 1).
SQ SEQUENCE 337 AA; 37290 MW; 183FE2F69B34CBF CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 337;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 69 IYEDDI 74

RESULT 7
YMG7_MYCTU
ID YMG7_MYCTU STANDARD; PRT; 388 AA.
AC Q50695;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2267C.
GN RV2267C OR MT2329 OR MYCY339.43.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Mycobacteriaceae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV:
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Teketa C.,
RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Stulton J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=CDL 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey L.,
RA Kolonay J.F., Nelson W.C., Umayan L.A., Emdolova M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; 277163; CAB00968.1; -
DR EMBL; AE007076; AAK46611.1; -
DR TIGR; MT2329; -
DR Tuberculist; RV2267c; -
DR InterPro: IPR000863; Sulfoltransferase.
DR Pfam; PF00685; Sulfoltransfer; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 388 AA; 46062 MW; 5DED0263275A9B24 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 388;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 302 LRYEDL 307

RESULT 8
MOEA_ANASP STANDARD: PRT; 436 AA.
AC 044243;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moea.
GN MOEA OR ALU5136.
OS Anabaena sp. (Strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=96272265; PubMed=8682795;
RA Ramasamy K.S., Endley S., Golden J.W.;
RT "Nitrate reductase activity and heterocyst suppression on nitrate in
RL Anabaena sp. strain PCC 7120 require moea."
J. Bacteriol. 178:3893-3898(1996).
[2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kurita T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Kuraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL Nucleic Acids Res. 29:205-213(2001).
CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF A DEMOLYBDO-COFACTOR
CC (MOLYBDOPEPTIN), NECESSARY FOR MOLYBDO-ENZYMES (BY SIMILARITY).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: TO THE C-TERMINAL OF CINNAMON/GERANYRIN.
-----
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; U34309; AAC44505.1; -
DR EMBL; AP003598; BAB76835.1; -
DR InterPro: IPR001453; MOCF_biosynth.
DR InterPro: IPR005111; Moea_C.
DR InterPro: IPR005110; Moea_N.
DR Pfam; PF00994; MOCF_biosynth; 1.
DR Pfam; PF03453; Moea_N; 1.
DR Pfam; PF03454; Moea_C; 1.
DR Pfam; PF002460; MOCF_biosynth; 1.
DR ProDom; PD002460; MOCF_biosynth; 1.
DR TIGR; TIGR00177; molyb_syn; 1.
DR PROSITE; PS01079; MOCF BIOSYNTHESIS 2; 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 436 AA; 47332 MW; 65D732E44336D561 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 436;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDV 89
Db 84 VRYEDV 89

RESULT 9
C6ST_CHICK STANDARD: PRT; 458 AA.
AC 092179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Chondroitin 6-sulfoltransferase (EC 2.8.2.17) (C6ST).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-44; 46-52 AND 186-192.
RC STRAIN=White Leghorn; TISSUE=Embryonic chondrocytes;
RX MEDLINE=95355490; PubMed=7629189;
RA Fukuta K., Uchimura K., Nakashima K., Kato M., Kimata K.,
RA Shinomura T., Habuchi O.;
RT "Molecular cloning and expression of chick chondrocyte chondroitin 6-
RT sulfoltransferase."
J. Biol. Chem. 270:18575-18580(1995).
RL
CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN
CC SULFATE.
CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + chondroitin =
CC adenosine 3',5'-bisphosphate + chondroitin 6'-sulfate.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE

```

(BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D49915; BAA08655.1; -
KM Transferrase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
FT DOMAIN
FT TRANSFER 1
FT TRANSFER 23
FT TRANSFER 24
FT TRANSFER 37
FT TRANSFER 38
FT TRANSFER 458
FT TRANSFER 62
FT TRANSFER 73
FT TRANSFER 95
FT TRANSFER 236
FT TRANSFER 399
FT TRANSFER 443
FT TRANSFER 458
SQ SEQUENCE 458 AA; 52253 MW; C9A3B7D0A5086F0C CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 458;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 348 VREEDV 353

RESULT 10
AHPE-STAAU STANDARD; PRT; 507 AA.
ID AHPE-STAAU STANDARD; PRT; 507 AA.
AC 005204;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALK1 hydroperoxide reductase subunit F (EC 1.6.4.-).
GN AHPE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RM4220.
RA Jones E.C., Francis K.P., Stewart G.S.A.B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SERVES TO PROTECT THE CELL AGAINST DNA DAMAGE BY ALKYL
CC HYDROPEROXIDES. IT CAN USE EITHER NADH OR NADPH AS ELECTRON DONOR
CC FOR DIRECT REDUCTION OF REDOX DYES OR OF ALKYL HYDROPEROXIDES WHEN
CC COMBINED WITH THE AHPC PROTEIN.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-II.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U92441; AAB51152.1; -
DR HSSP: P09625; 1TRB.
DR InterPro: IPR001327; FAD_Pyr_redox.
DR InterPro: IPR000103; Pyridine_redox_2.
DR Pfam: PF00070; Pyr_redox; 1.
DR PRINTS: PR00368; FADPNR.

DR PRODOM: PD000139; FAD_Pyr_redox; 1.
DR PROSITE: PS00573; PYRIDINE_REDOX_2; 1.
KM Redox-active center; Oxidoreductase; NADP; NAD; Flavoprotein; FAD.
FT NP_BIND 207 222
FT DISULFID 335 338
FT NP_BIND 347 361
FT NP_BIND 467 477
FT NP_BIND 467 477
SQ SEQUENCE 507 AA; 54720 MW; EDED04232731D21E CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 507;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 416 IREEDM 421

RESULT 11
M3K3-MOUSE STANDARD; PRT; 626 AA.
ID M3K3-MOUSE STANDARD; PRT; 626 AA.
AC 061084;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.-)
DE (MAPK/ERK kinase kinase 3) (MEK kinase 3) (MEK 3).
GN MAPK3 OR MEK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RA MEDLINE=96214986; PubMed=8621389.
RA Blank J.L., Gervais P., Elliott E.M., Sather S., Johnson G.L.;
RT "Molecular cloning of mitogen-activated protein kinase kinases
RT (MEK) 2 and 3. Regulation of sequential phosphorylation pathways
RT involving mitogen-activated protein kinase and c-Jun kinase.";
RL J. Biol. Chem. 271:5361-5368(1996).
CC -1- FUNCTION: PREFERENTIALLY ACTIVATES P42/44 (ERK2/ERK) MAP KINASES.
CC -1- SIMILARITY: BELONGS TO THE SPR/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U43187; AAB03535.1; -
DR HSSP: Q00534; 1B18.
DR MGD: MGI:1346874; MapK3.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_Thr_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR PRODOM: PD000001; Euk_Pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 362 622
FT NP_BIND 368 375
FT BINDING 391 391
FT ACT_SITE 489 489
SQ SEQUENCE 626 AA; 70775 MW; 00EF244C9556EDB CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 626;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYEDL 6
|||||
Db 64 VREYEDV 69

RESULT 12

YJEP-ECOLI STANDARD; PRT: 1107 AA.

AC P39285; P76798; Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Hypothetical protein yjep precursor.

GN YJEP OR B4159.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

CC Escherichia.

OX NCBI_TaxID=562;

RM SEQUENCE FROM N.A.

RP STRAIN-K12 / MG1655;

RC MEDLINE=97426617; PubMed=9278503;

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RM [3]

RP SEQUENCE OF 1-27 FROM N.A.

RC SPMRAIN-K12;

RX MEDLINE=88298809; PubMed=3042771;

RA Li Q.X., Dowhan W.;

RT "Structural characterization of Escherichia coli phosphatidylserine

RT decarboxylase.";

RL J. Biol. Chem. 263:11516-11522(1988).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE OPG003 FAMILY. STRONG, TO H. INFLUENZA

CC H10195.1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U14003; AAA97058.1; -

DR EMBL: AE000488; AAC77119.1; -

DR EMBL: J03916; AAA83897.1; ALT_INT.

DR EcoGene; EG12478; yjep

DR InterPro; IPR001880; MSion_channel.

DR Pfam; PF00924; MS_channel.1.

DR PROSITE; PS01246; UPF0003.1.

KW Hypothetical protein; Transmembrane; Signal; Complete proteome.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 1107 HYPOTHELTICAL PROTEIN YJEP.

FT TRANSMEM 467 487 POTENTIAL.

FT TRANSMEM 551 571 POTENTIAL.

FT TRANSMEM 600 620 POTENTIAL.

FT TRANSMEM 628 648 POTENTIAL.

FT TRANSMEM 674 694 POTENTIAL.

FT TRANSMEM 698 718 POTENTIAL.

FT TRANSMEM 785 805 POTENTIAL.

FT TRANSMEM 828 848 POTENTIAL.

FT TRANSMEM 875 895 POTENTIAL.

FT TRANSMEM 910 930 POTENTIAL.

FT TRANSMEM 930 950 POTENTIAL.

FT CONFLICT 1015 1015 R -> A (IN REF. 1).

FT CONFLICT 1015 1015 R -> A (IN REF. 1).

SQ SEQUENCE 1107 AA; 123967 MW; 5F52A2993B90532B CRC64;

Query Match 90.3%; Score 28; DB 1; Length 1107;

Best Local Similarity 83.3%; Pred. No. 2.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYEDL 6

Db 348 VREYEDL 353

RESULT 13

RRPL_P13H4 STANDARD; PRT: 2233 AA.

AC P12577;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE RNA polymerase beta subunit (Ec 2.7.7.48) (large structural protein)

DE (L. protein).

GN L.

OS Human parainfluenza 3 virus (strain NIH 47885).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Respirovirus.

OX NCBI_TaxID=11217;

RM [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88306242; PubMed=2841798;

RA Gallinski M.S., Mink M.A., Pons M.W.;

RT "Molecular cloning and sequence analysis of the human parainfluenza 3

RT virus gene encoding the L protein.";

RL Virology 165:499-510(1988).

RM [2]

RP SEQUENCE OF 1-26 FROM N.A.

RX MEDLINE=88032139; PubMed=2822598;

RA Storey D.G., Cote M.-J., Dimock K., Kang C.Y.;

RT "Nucleotide sequence of the coding and flanking regions of the human

RT parainfluenza virus 3 hemagglutinin-neuraminidase gene: comparison

RT with other paramyxoviruses.";

RL Intervirology 27:69-80(1987).

CC -1- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY

CC FUNCTION IN RNA SYNTHESIS, CAPING, METHYLATION AND POLY(A)

CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAS. RNA EDITING OF THE P

CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC {RNA}(N).

CC {RNA}(N).

CC -1- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M21649; AAA46854.1; -

DR EMBL: M20402; AAA46857.1; -

DR PIR: A29246; ZLN2P3.

DR InterPro; IPR001016; Viral_RNA_pol_L.

DR Pfam; PF00946; Paramyx_RNA_pol.1.

KW Transferase; RNA-directed RNA polymerase.

FT CONFLICT 26 26 K -> R (IN REF. 2).

SQ SEQUENCE 2233 AA; 255802 MW; DBB8B5DE74B4638 CRC64;

Query Match

90.3%; Score 28; DB 1; Length 2233;

Query Match	87.1%;	Score 27;	DB 1;	Length 157;
Best Local Similarity	83.3%;	Pred. No. 42;		

```

RL      gene 114:113-119(1992):
CC      -1- FUNCTION: ESSENTIAL FOR DNA SYNTHESIS
CC      -----
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M82842; AAA32597.1; -; 98442FCB05FA7D CRC64;
SQ SEQUENCE 225 AA; 25858 MW;

Query Match 87.1%; Score 27; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
Db 32 VRYED 36

RESULT 17

YUBP_ECOLI

ID YUBP_ECOLI STANDARD; PRT; 273 AA.

AC P18005; P52149;

DT 01-NOV-1990 (Rel. 16, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein yubp.

GN YUBP OR YGBB.

OS Escherichia coli.

OC Plasmid F, and Plasmid IncFII R1.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC PLASMID-F;

RX MEDLINE=90136505; PubMed=2693941;

RA Loh S., Gram D., Skurray R.;

RT "Nucleotide sequence of the leading region adjacent to the origin of
RT transfer on plasmid F and its conservation among conjugative
RT plasmids.";

RL Mol. Genet. 219:177-186(1989).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / CR63; PLASMID-F;

RA Shimizu H., Satoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;

RT "Complete nucleotide sequence of the F plasmid: its implications for
RT organization and diversification of plasmid genomes.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

FT VARIANT 210 210 W -> R (IN PLASMID INCFII R1).
SQ CONFLICT 34 34 N -> T (IN REF. 1).
SQ SEQUENCE 273 AA; 31615 MW; F136DC9182E80CFA CRC64;

Query Match 87.1%; Score 27; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
Db 189 VRYED 193

RESULT 18

GDFP_HUMAN

ID GDFP_HUMAN STANDARD; PRT; 308 AA.

AC Q99868; P78360; Q14629; Q9NRT0;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Growth/differentiation factor 15 precursor (GDF-15) (Placental bone
DE morphogenic protein) (Placental TGF-beta) (Macrophage inhibitory
DE cytokine-1) (MIF-1) (Prostate differentiation factor) (NSAID-regulated
DE protein 1) (NRG-1).

GN GDF15 OR PLAB OR PTGFB OR MIF1 OR PDF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Fibrosarcoma;

RX MEDLINE=98006316; PubMed=9348093;

RA Yokoyama-Kobayashi M., Seeki M., Sekine S., Kato S.;

RT "Human cDNA encoding a novel TGF-beta superfamily protein highly
RT expressed in placenta.";

RL J. Biochem. 122:622-626(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=97470998; PubMed=9326641;

RA Baotcov M.R., Bauskin A.R., Valenzuela S.M., Moore A.G., Bansal M.,
RA He X.Y., Zhang H.P., Donnellan M., Mahler S., Pryor K., Walsh B.J.,
RA Nicholson R.C., Fairlie W.D., Por S.B., Robins J.M., Breit S.N.;RT "MIC-1, a novel macrophage inhibitory cytokine, is a divergent member
RT of the TGF-beta superfamily.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:11514-11519(1997).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98041637; PubMed=9375789;

RA Hromas R., Huford M., Sutton J., Xu D., Li Y., Lu L.;

RT "PLAB, a novel placental bone morphogenetic protein.";

RL Biochim. Biophys. Acta 1354:40-44(1997).

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98256302; PubMed=9593718;

RA Parakekar V.M., Vail A.L., Grasser W.A., Brown T.A., Xu H.,
RA Vukilevic S., Ke H.Z., Qi H., Owen T.A., Thompson D.D.;RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";

RL J. Biol. Chem. 273:13760-13767(1998).

RN [5]

RP SEQUENCE OF 14-308 FROM N.A.

RX MEDLINE=98085971; PubMed=9426002;

RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Qiu L., Baumes S.A.,
RA Marcellino R.A., de Jesus G.M., Wellington S., Knowles J.A.,
RA Warburton D., Brown S., Soares M.B.;RT "Identification of a novel member of the TGF-beta superfamily highly
RT expressed in human placenta.";

RL Gene 203:17-26(1997).

RN [6]

RP SEQUENCE OF 264-308 FROM N.A.

RA Baek S.J., Nixon J., Eling T.;
 RT "NRG-1 is associated with apoptosis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBMITTER: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, WITH LOWER
 CC LEVELS IN PROSTATE AND COLON AND SOME EXPRESSION IN KIDNEY.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB000584; BAA19151.1; -
 DR EMBL: AF019770; AAB88673.1; -
 DR EMBL: U88323; AAB88913.1; -
 DR EMBL: AF003934; AAC24456.1; -
 DR EMBL: AF008303; AAC39537.1; -
 DR EMBL: AF173860; AAF89834.1; -
 DR HSSP: P18075; IAMP.
 DR MIM: 605312; -
 DR InterPro: IPR002400; GFCysknot.
 DR InterPro: IPR001839; TGFb.
 DR Pfam: PF00019; TGF-Beta; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR ProDom: PD000357; TGFb; 1.
 DR SMART: SM00204; TGFb; 1.
 DR PROSITE: PS00250; TGF-BETA.1; FALSE NEG.
 KW Growth factor; Cytokine; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 29
 FT PROPEL 30 194
 FT CHAIN 195 308
 FT DISULFID 211 274
 FT DISULFID 240 305
 FT DISULFID 244 307
 FT DISULFID 273 273
 FT CARBOHD 70 78
 FT VARIANT 48 48
 FT CONFLICT 9 9
 FT CONFLICT 202 202
 FT CONFLICT 269 269
 FT CONFLICT 288 288
 FT SEQUENCE 308 AA; 34168 MW; ADF3A3ED065ACAZE CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 308;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RYEDL 6
 DB 59 RYEDL 63
 RESULT 19
 MRAM_BACSU
 ID MRAM_BACSU STANDARD; PRT; 311 AA.
 AC 007876; 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE S-adenosyl-methyltransferase mram (EC 2.1.1.-).
 GN MRAM.
 OS Bacillus subtilis.
 OC Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1423;
 RX NCBI_TaxID=1423;
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;

RA Daniel R.A., Williams A.M., Errington J.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 108-311 FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=94064553; PubMed=8244929;
 RX Yanouri A., Daniel R.A., Errington J., Buchanan C.E.;
 RT "Cloning and sequencing of the cell division gene pphB, which encodes
 RT pilingin-binding protein 2B in Bacillus subtilis."
 RL J. Bacteriol. 175:7604-7616(1993).
 CC -1- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z68230; CAA92525.1; -
 DR EMBL: L09703; AAC36835.1; -
 DR EMBL: Z59111; CAB13587.1; -
 DR Subtilist; BG10219; mram.
 DR InterPro: IPR002903; Bac_Metrifase.
 DR Pfam: PF01795; Methyltransf.5; 1.
 DR ProDom: PD004685; Bac_Metrifase; 1.
 DR TIGRFAWS: TIGR00006; UPE0117.1
 KW Transferase; Methyltransferase; Complete proteome.
 SEQUENCE 311 AA; 35305 MW; C50C609129DFDAD CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RYEDL 6
 DB 143 RYEDL 147
 RESULT 20
 REBB_NEIMA
 ID REBB_NEIMA STANDARD; PRT; 341 AA.
 AC 095642; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE drpD-glucose 4,6-dehydratase (EC 4.2.1.46).
 GN (REBB) OR NMA0189) AND (REBB2 OR NMA0204).
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-1 / Serogroup A;
 RX MEDLINE=99150278; PubMed=10024588;
 RA Lee F.K., Gibson B.W., McLaugh W., Zaleski A., Apicella M.A.;
 RT "Relationship between UDP-glucose 4-epimerase activity and
 RT oligoglucose glycoforms in two strains of Neisseria meningitidis."
 RL Infect. Immun. 67:1405-1414(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Raftery M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria

RT meningitidis 22491.1;
 RL Nature 404:502-506(2000).
 CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
 CC glucose + H(2)O.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: DTD-P-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
 CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTD-P-GLUCOSE
 CC DEHYDRATASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF083467; AAD23919.1; -;
 DR EMBL: AL162752; CAB83504.1; -;
 DR EMBL: AL162752; CAB83518.1; -;
 DR HSSP: P27830; IBXK.
 DR InterPro: IPR001509; Epimerase_Dh.
 DR Pfam: PF01370; Epimerase; 1.
 DR TIGRfams: TIGR01181; dTDP-gluc_dehyd; 1.
 DR Lipopolysaccharide biosynthesis; Lyase; NAD; Complete proteome.
 KW NP_BIND 8 NAD (POTENTIAL).
 FT SEQUENCE 341 AA; 38224 MW; 826023B63916E72A CRC64;
 SQ
 Query Match 87.1%; Score 27; DB 1; Length 341;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 RYEDL 6
 Db 283 RYEDL 287
 RESULT 21
 REFB_NEIGO STANDARD; PRT; 346 AA.
 ID P37761;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE dTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
 GN REFB.
 OS Neisseria gonorrhoeae.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 CX NCBI_TaxID=485;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MS11;
 RX MEDLINE=95050260; PubMed=7961452;
 RA Robertson B.D., Frosch M., van Putten J.P.M.;
 RA "The identification of cryptic rhamnose biosynthesis genes in
 RA Neisseria gonorrhoeae and their relationship to lipopolysaccharide
 RT biosynthesis.";
 RL J. Bacteriol. 176:6915-6920(1994).
 CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
 CC glucose + H(2)O.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: DTD-P-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
 CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTD-P-GLUCOSE
 CC DEHYDRATASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: 232742; CAB83652.1; -;
 DR EMBL: 221508; CAB79718.1; -;
 DR PIR: S47045; S47045.
 DR HSSP: P27830; IBXK.
 DR InterPro: IPR001509; Epimerase_Dh.
 DR Pfam: PF01370; Epimerase; 1.
 DR TIGRfams: TIGR01181; dTDP-gluc_dehyd; 1.
 KW Lipopolysaccharide biosynthesis; Lyase; NAD.
 FT NP_BIND 13 19 NAD (POTENTIAL).
 SQ SEQUENCE 346 AA; 38828 MW; B52FD8125C567675 CRC64;
 Oy 2 RYEDL 6
 Db 288 RYEDL 292
 RESULT 22
 REFB_NEIMB STANDARD; PRT; 355 AA.
 ID P55294; Q9J514;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE dTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
 GN (REFBI OR NMB0063) AND (REFB2 OR NMB0079).
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 CX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B1940 / Serogroup B;
 RX MEDLINE=944293762; PubMed=8022285;
 RA Hammerschmidt S., Birkholz C., Zahnlinger U., Robertson B.D.,
 RA van Putten J.P.M., Ebeling O., Frosch M.;
 RA "Contribution of genes from the capsule gene complex (cps) to
 RT lipopolysaccharide biosynthesis and serum resistance in Neisseria
 RT meningitidis.";
 RL Mol. Microbiol. 11:885-896(1994).
 CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
 CC glucose + H(2)O.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: DTD-P-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
 CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTD-P-GLUCOSE
 CC DEHYDRATASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: L09188; AAA63157.1; -.
DR EMBL: AE002365; AAP40531.1; -.
DR EMBL: AE002367; AAP40543.1; -.
DR HSSP: P27830; IBXK.
DR TIGR: NME0063; -.
DR TIGR: NME0079; -.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
DR TIGRFAMs: TIGR01181; drdp_gluo_dehyd; 1.
KM Lipopolysaccharide biosynthesis; Lyase; NAD; Complete proteome.
FT NP_LIND 8 14 NAD (POTENTIAL).
FT CONFLICT 1 3 MKK -> MQTAKKT (IN REF. 1).
FT CONFLICT 26 28 RDA -> ODS (IN REF. 1).
FT CONFLICT 32 32 V -> L (IN REF. 1).
FT CONFLICT 46 47 EV -> DI (IN REF. 1).
FT CONFLICT 73 73 Y -> H (IN REF. 1).
FT CONFLICT 142 142 G -> H (IN REF. 1).
FT CONFLICT 152 152 A -> T (IN REF. 1).
FT CONFLICT 268 268 A -> T (IN REF. 1).
FT CONFLICT 274 274 A -> V (IN REF. 1).
SQ SEQUENCE 355 AA; 39865 MW; 631AAOEDA02B6F41 CRC64;

Query Match
Best Local Similarity 87.1%; Score 27; DB 1; Length 355;
Matches 5; Conservative 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 283 RYEDL 287

RESULT 23
PK_STAAM STANDARD: PRT; 396 AA.
AC Q9ZSC4;
DI 30-MAY-2000 (Rel. 39, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR SAV0773 OR SA0728 OR MM0735.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=138876, 158679, 196620, 1280;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RC MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyama A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Hattori M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RT Lancet 357:1225-1240(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RC MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RA "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RT Lancet 359:1819-1827(2002).
[3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=BB;
RA Morrissey J.A., Williams P.;
RT "Isolation and characterisation of a glycolytic operon in
RT Staphylococcus aureus";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -1 PATHWAY: Second phase of glycolysis; second step.
CC -1 SUBUNIT: MONOMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
CC -1 SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.sib-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: AP003360; BAB56935.1; -.
DR EMBL: AP003131; BAB41961.1; -.
DR EMBL: AP004824; BAB4600.1; -.
DR EMBL: AJ133520; CAB38646.1; -.
DR HSSP: P36204; IVPE.
DR InterPro: IPR001576; PGK.
DR Pfam: PF00162; PGK; 1.
DR PRINTS: PRO0477; PHGLYKINASE.
DR PROSITE: PS00111; PGLYCERATE_KINASE; 1.
KM Transferrase; kinase; glycolysis; Complete proteome.
SQ SEQUENCE 396 AA; 42601 MW; C8383CFB5BF43C CRC64;

Query Match
Best Local Similarity 87.1%; Score 27; DB 1; Length 396;
Matches 5; Conservative 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 24
LE21_ARCFU STANDARD: PRT; 418 AA.
ID LE21_ARCFU
AC O28316;
DI 15-JUL-1998 (Rel. 36, Created)
DI 15-JUN-2002 (Rel. 41, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-Isopropylmalate dehydratase large subunit 1 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 1) (Alpha-IRM isomerase 1) (IPMI 1).
GN LE21 OR AF1963.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RC MEDLINE=98045343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus";

```

RL Nature 390:364-370(1997).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000967; AAB89290.1; ALT_INT.
DR TIGR: AF1963;
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF00330; aconitase_1.
DR ProDom: PD000511; Aconitase_N; 1.
DR PROSITE: PS00450; ACONITASE_1; 1.
DR PROSITE: PS01244; ACONITASE_2; 1.
KM Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 418 AA; 45140 MW; 51B2BD0A3D9CA22 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 302 RYEDL 306
|||||
LE22_METRA STANDARD; PRT; 418 AA.
ID LE22_METRA
AC ORFV2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydratase large subunit 1 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 1) (Alpha-IPM isomerase 1) (IPM1 1).
GN LEUC1 OR MK1440.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
CC Methanopyrus.
CC NCBI_TaxID=2320;
CX
RN
RM
SEQUENCE FROM N.A.
RX STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; Pubmed=11930014;
RA Shesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rozozin I.B., Telusov R.L., Wolf Y.I., Stettin K.O.,
RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE010436; AM02653.1;
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF00330; aconitase_1.
DR ProDom: PD000511; Aconitase_N; 1.
DR PROSITE: PS00450; ACONITASE_1; 1.
DR PROSITE: PS01244; ACONITASE_2; FALSE NEG.
KM Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 418 AA; 45202 MW; FABA735A9BD51C72 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 302 RYEDL 306
|||||
LE22_METTH STANDARD; PRT; 428 AA.
ID LE22_METTH
AC O27668;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydratase large subunit 2 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 2) (Alpha-IPM isomerase 2) (IPM1 2).
GN LEUC2 OR MTH151.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
CC Methanobacteriaceae; Methanothermobacter.
CC NCBI_TaxID=187420;
CX
RN
RM
SEQUENCE FROM N.A.
RX STRAIN-Delta H;
RX MEDLINE=98037514; Pubmed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Pradhan S.,
RA McDougall S., Shlmer C., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----

CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE010436; AM02653.1;
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF00330; aconitase_1.
DR ProDom: PD000511; Aconitase_N; 1.
DR PROSITE: PS00450; ACONITASE_1; 1.
DR PROSITE: PS01244; ACONITASE_2; FALSE NEG.
KM Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 418 AA; 45202 MW; FABA735A9BD51C72 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 302 RYEDL 306
|||||
LE22_METTH STANDARD; PRT; 428 AA.
ID LE22_METTH
AC O27668;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydratase large subunit 2 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 2) (Alpha-IPM isomerase 2) (IPM1 2).
GN LEUC2 OR MTH151.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
CC Methanobacteriaceae; Methanothermobacter.
CC NCBI_TaxID=187420;
CX
RN
RM
SEQUENCE FROM N.A.
RX STRAIN-Delta H;
RX MEDLINE=98037514; Pubmed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Pradhan S.,
RA McDougall S., Shlmer C., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE000922; AAB6104.1; -
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF000330; aconitase; 1.
DR PROSITE; PS00450; ACONITASE_1; 1.
DR PROSITE; PS01244; ACONITASE_2; 1.
KW Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 304 304 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 364 364 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 367 367 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 428 AA; 46447 MW; B93FIDC53C9B8178 CRC64.

Query Match 87.1%; Score 27; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 308 RYEDL 312

RESULT 27
DCDA_BACHD STANDARD; PRT: 439 AA.
AC O9KCM5:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Diaminopimelate decarboxylase (EC 4.1.1.20) (DAP decarboxylase).
GN LYSA OR BH1544.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Mieno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogatawara N., Kunara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: Meso-2,6-diaminoheptanedioate = L-lysine +
CC CO(2).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Lysine biosynthesis; last step.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
CC DECARBOXYLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AP001512; BAB05263.1; -
DR InterPro: IPR000183; Decarboxylase2.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
DR Pfam; PF02784; Orn_Arg_dec; 1.
DR PRINTS; PR01179; OADACRXLASE.
DR TIGRFAMs; TIGR01048; lysA; 1.
DR PROSITE; PS00878; ODR_DC_2.1; FALSE_NEG.

DR PROSITE; PS00879; ODR_DC_2.2; 1.
KW Lysine biosynthesis; Lyase; Decarboxylase; Pyridoxal phosphate;
KW Complete proteome.
FT BINDING 66 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 439 AA; 48444 MW; FE59E21033E6E90F CRC64.

Query Match 87.1%; Score 27; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 422 RYEDL 426

RESULT 28
GUAD_ECOLI STANDARD; PRT: 439 AA.
ID GUAD_ECOLI
AC P76641; Q46816;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine deaminase (EC 3.5.4.3) (Guanase) (Guanine aminase) (Guanine
DE aminohydrolase) (GAH).
GN GUAD OR B2883.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655.
RX MEDLINE=9742617; PubMed=9278503.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=20372655; PubMed=10913105;
RA Maynes J.T., Yuan R.G., Snyder F.F.;
RT "Identification, expression, and characterization of Escherichia coli
RT guanine deaminase."
RL J. Bacteriol. 182:4658-4660(2000).
CC -1- FUNCTION: CATALYZES THE HYDROLYTIC DEAMINATION OF GUANINE,
CC PRODUCING XANTHINE AND AMMONIA.
CC -1- CATALYTIC ACTIVITY: Guanine + H(2)O = xanthine + NH(3).
CC -1- COFACTOR: CONTAINS 1 MOLE OF ZINC PER SUBUNIT.
CC -1- PATHWAY: PURINE NUCLEOTIDE CATABOLISM.
CC -1- SIMILARITY: BELONGS TO THE ATZ/TRZ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE000372; NAC75921.1; -
DR EMBL; U28375; AAA83064.1; -
DR EcGene; EGI3066; guad.
DR InterPro: IPR002604; ATZ_TRZ.
DR Pfam; PF01685; ATZ_TRZ; 1.
KW Hydrolyase; Zinc; Complete proteome.
SQ SEQUENCE 439 AA; 50244 MW; 43389F3AF9E4AD83 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
 DB 112 RYEDL 116

RESULT 29
 ID C133_DROME STANDARD; PRT; 492 AA.
 AC Q9VGB3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cytochrome P450 313a3 (EC 1.14.-.-) (CYPCXCXIII1A3).
 GN CYP313A3 OR CG10093.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Jorgensen C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kienison J.A., Kiechum K.A.,
 RA Kimmel B.E., Kodira C.D., Kravitz S., Kulp D., Lai Z., Liang Y., Lin X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Nelson B.;
 RL Unpublished observations (SEP-2000).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
 CC IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) -> ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
 CC (potential).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous

CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.slb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: AE003695; AAF54770.2; ALR_SEQ.
 CC F1yase: F8n0038007; CYP313a3.
 CC InterPro: IPR001128; Cytochrome_P450.
 CC Pfam: PF00067; P450.1.
 CC PRINTS: PR00385; P450.
 CC PROSITE: PS00086; CYTOCHROME_P450.1.
 CC OXIDOREDUCTASE: Monooxygenase; Membrane; Heme; Microsome;
 CC Endoplasmic reticulum; Hypothetical protein.
 CC BINDING 438 438 HEME (BY SIMILARITY).
 CC SEQUENCE 492 AA; 56195 MW; 2836DJ3953C3B4DC CRC64;
 SO

Query Match 87.1%; Score 27; DB 1; Length 492;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
 DB 466 RYEDL 470

RESULT 30
 ID FUT3_ARATH STANDARD; PRT; 493 AA.
 AC Q9CA71;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable fucosyltransferase 3 (EC 2.4.1.-) (ATFUT3).
 GN FUT3 OR ATIG74420 OR FIM20.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhahn T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizlar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.-J., Koo H.L., Kremenetskaia I., Kutz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marshall A.,
 RA Miltner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Souhvik A.M.,
 RA Sun H., Talon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*


```

RT thaljana."
RL Nature 408:816-820(2000).
CC -1- FUNCTION: May be involved in cell wall biosynthesis. May
CC act as a fucosyltransferase.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
CC form in trans cisternae of Golgi (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in roots, stems, leaves, flowers,
CC siliques and seedlings.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 37.
CC -1- CAUTION: THE SEQUENCE IN REF.2 DIFFERS FROM THAT SHOWN DUE TO
CC WRONG EXON PREDICTIONS FROM THE GENOMIC SEQUENCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF417473; AAL50622.1; -.
DR EMBL: AC011765; AAG52352.1; ALU_SEQ.
DR InterPro: IPR004938; XG_Ftase.
DR Pfam: PF03254; XG_Ftase; 1.
KW Transferase; glycosyltransferase; Transmembrane; glycoprotein;
KW Signal-anchor; Golgi stack; Cell wall.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 32 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 33 493 LEMNENL, CATALYTIC (POTENTIAL).
FT DOMAIN 376 379 POLY-SER.
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 493 AA; 55997 MW; 64CB98C78C26DEA3 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 108 RYEDL 112

RESULT 31
Y213_HAEIN
ID Y213_HAEIN STANDARD; PRT; 514 AA.
AC P44572;
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Putative binding protein H10213 precursor.
GN H10213.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utechtback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae

```

```

RT Rd";
RL Science 269:496-512(1995).
CC -1- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32706; AAC21881.1; -.
DR HSPG: P06202; IBS2.
DR TIGR: H10213; -.
DR InterPro: IPR000437; Prok_Lipoprot.
DR InterPro: IPR000914; ssp_bac_5.
DR Pfam: PF00496; ssp_bac_5; 1.
DR PROSITE: PS01040; SSP_BACTERIAL_5; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; FALSE_NEG.
KW Hypothetical protein; Transport; Membrane; Signal; Lipoprotein;
KW Complete proteome.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 514 PUTATIVE BINDING PROTEIN H10213.
FT LIPID 24 24 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 514 AA; 58876 MW; 789188C4328BDEBC CRC64;

Query Match 87.1%; Score 27; DB 1; Length 514;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 359 IRYEDL 364

RESULT 32
PPCK_PYRO
ID PPCK_PYRO STANDARD; PRT; 621 AA.
AC O58050;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32) (PPP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN PCKG OR PHO312.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxId=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=96344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Kikuchi H.;
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2);
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]

```

```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AP000001; BAA29385.1; -
CC      InterPro: IPR000364; PEP_carboxykin.
CC      Pfam: PF00821; PEPCK: 1.
CC      ProDom: PD004738; PEP_carboxykin; 1.
CC      PROSITE: PS00505; PEPCK_GTP; FALSE_NEG.
CC      Glucocoenogenesis; Lyase; Decarboxylase; GTP-binding; Complete proteome.
CC      NP_BIND 218 225 GTP (POTENTIAL).
CC      ACT_SITE 269 269 BY SIMILARITY.
CC      SEQUENCE 621 AA; 72201 MW; 738232FE30F20F01 CRC64;
CC
CC      Query Match      87.1%; Score 27; DB 1; Length 621;
CC      Best Local Similarity 100.0%; Pred. No. 1.9e+02;
CC      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CY      2 RYEDL 6
DB      537 RYEDL 541

```

```

CC      RESULT 33
CC      SKN1_CABEL STANDARD; PRT; 623 AA.
CC      AC P34707;
CC      DT 01-FEB-1994 (Rel. 28, Created)
CC      DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC      DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC      DE Skn-1 protein.
CC      SKN-1 OR T19E7.2.
CC      Caenorhabditis elegans.
CC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
CC      OC Rhabdilitidae; Peloderinae; Caenorhabditis.
CC      OX NCBI_TaxID=6239;
CC      [1]
CC      RN SEQUENCE FROM N.A.
CC      RC STRAIN-Bristol N2;
CC      RA Pauley A., Gattung S.;
CC      RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC      RN [2]
CC      RP REVISIONS.
CC      RA Waterston R.;
CC      RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC      RN [3]
CC      RP SEQUENCE OF 49-623 FROM N.A., FUNCTION, AND DEVELOPMENTAL STAGE.
CC      RC STRAIN-Bristol N2;
CC      RX MEDLINE=92191285; PubMed=1547503;
CC      RA Bowerman B., Eaton B.A., Priess J.R.;
CC      RL "Skn-1, a maternally expressed gene required to specify the fate of
CC      ventral blastomeres in the early C. elegans embryo.";
CC      RL Cell 68:1061-1075(1992).
CC      RN [4]
CC      RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 546-619.
CC      RX MEDLINE=98290451; PubMed=9628487;
CC      RA Rupert P.B., Daughdrill G.W., Bowerman B., Matthews B.W.;
CC      RL "A new DNA-binding motif in the Skn-1 binding domain-DNA complex.";
CC      RL Nat. Struct. Biol. 5:484-491(1998).
CC      CC -1- FUNCTION: REQUIRED TO SPECIFY THE FATE OF VENTRAL BLASTOMERES IN
CC      THE EARLY EMBRYO, AND POSTEMBRYONICALLY FOR THE DEVELOPMENT OF THE
CC      INTESTINE.
CC      CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC      CC -1- DEVELOPMENTAL STAGE: EARLY EMBRYOGENESIS, AND POSTEMBRYONIC
CC      INTESTINAL CELLS. EXPRESSED MATERNALLY.
CC      CC -1- SIMILARITY: BELONGS TO THE SKN1 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U42843; AAA83594.2; -
CC      DR EMBL: M64359; -; NOT_ANNOTATED_CDS.
CC      DR PIR: A42143; A42143.
CC      DR PDB: 1SKN; 17-JUN-98.
CC      DR TRANSFAC: T01614; -
CC      DR WormRep: T19E7.2; CE27591.
CC      DR InterPro: IPR004827; TF_bZIP.
CC      DR PROSITE: PS00036; BZIP_BASIC; 1.
CC      KW Developmental protein; DNA-binding; Nuclear protein; 3D-structure.
CC      FT DNA_BIND 597 612 BASIC MOTIF.
CC      FT SEQUENCE 623 AA; 70708 MW; 31A7030AB7CC691 CRC64;
CC
CC      Query Match      87.1%; Score 27; DB 1; Length 623;
CC      Best Local Similarity 100.0%; Pred. No. 1.9e+02;
CC      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CY      2 RYEDL 6
DB      240 RYEDL 244

```

```

CC      RESULT 34
CC      PCK_PYRFU STANDARD; PRT; 624 AA.
CC      ID PCK_PYRFU
CC      AC Q80410;
CC      DT 15-JUN-2002 (Rel. 41, Created)
CC      DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC      DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC      DE Phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32) (PEP
CC      carboxylase) (Phosphoenolpyruvate carboxylase) (PEPCK).
CC      PKG OR P70283.
CC      GN Pyrococcus furiosus.
CC      OS Pyrococcus furiosus.
CC      OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC      OC Pyrococcus.
CC      OX NCBI_TaxID=2261;
CC      [1]
CC      RN SEQUENCE FROM N.A.
CC      RC STRAIN-Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
CC      RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
CC      RL "The complete sequence of the Pyrococcus furiosus genome.";
CC      RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC      CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC      + CO(2).
CC      CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC      CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AE010153; AAL80413.1; -
CC      DR PROSITE: PS00505; PEPCK_GTP; FALSE_NEG.
CC      KW Glucocoenogenesis; Lyase; Decarboxylase; GTP-binding; Complete proteome.
CC      FT NP_BIND 223 230 GTP (POTENTIAL).
CC      FT ACT_SITE 274 274 BY SIMILARITY.
CC      SEQUENCE 624 AA; 72665 MW; EC6D925CF57EAC25 CRC64;
CC
CC      Query Match      87.1%; Score 27; DB 1; Length 624;
CC      Best Local Similarity 100.0%; Pred. No. 1.9e+02;

```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 542 RYEDL 546

RESULT 35
GIDA_BUCAI
ID GIDA_BUCAI STANDARD: PRT; 628 AA.
AC P57117;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Glucose inhibited division protein A.
GN GIDA OR BU001.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001118; BAB12729.1; -
CC InterPro: IPR002218; GIDA.
CC InterPro: IPR004416; GIDA_sub.
CC Pfam: PF01134; GIDA.1.
CC Pfam: PF003738; GIDA.1.
CC ProDom: PD003738; GIDA.1.
CC TIGRfam: TIGR00136; gida.1.
CC PROSITE; PS01280; GIDA_1; 1.
CC PROSITE; PS01281; GIDA_2; 1.
CC Complete proteome.
CC SEQUENCE 628 AA; 70307 MW; 0653745332BC077B CRC64;

Query Match 87.1%; Score 27; DB 1; Length 628;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RYEDL 6
DB 518 RYEDL 523

RESULT 36
PEEL_HUMAN
ID PEEL_HUMAN STANDARD: PRT; 653 AA.
AC O14829; O15253; O9UT00; O9UT21;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Serine/threonine protein phosphatase with EF-hands-1 (EC 3.1.3.16)
DE (PPEF-1) (Protein phosphatase with EF calcium-binding domain) (PPEF)
DE (Serine/threonine protein phosphatase 7) (PP7).
GN PPEP1 OR PPEF OR PP7C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DB NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Retina;
RX MEDLINE=97471020; PubMed=9326663;
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,
RA Nathans J.;
RT "Identification and characterization of a conserved family of protein
serine/threonine phosphatases homologous to Drosophila retinal
degeneration C."
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE-Fetal brain;
RX MEDLINE=97358589; PubMed=9215685;
RA Montini E., Rugaili E.I., van de Vosse E., Andolfi G., Mariani M.,
RA Puccia A.A., Consales G.G., den Dunnen J.T., Ballabio A., Franco B.;
RT "A novel human serine-threonine phosphatase related to the Drosophila
retinal degeneration C (rdgc) gene is selectively expressed in sensory
neurons of neural crest origin."
RL Hum. Mol. Genet. 6:1137-1145(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Retina;
RX MEDLINE=98104127; PubMed=9430683;
RA Huang X., Honkanen R.E.;
RT "Molecular cloning, expression, and characterization of a novel human
serine/threonine protein phosphatase, PP7, that is homologous to
Drosophila retinal degeneration C gene product (rdgc).";
RL J. Biol. Chem. 273:1462-1468(1998).
RN [4]
RP SEQUENCE OF 1-355 FROM N.A.
RA Grahm D.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 356-653 FROM N.A.
RA Wray P.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY HAVE A ROLE IN THE RECOVERY OR ADAPTATION RESPONSE
CC OF PHOTORECEPTORS. MAY HAVE A ROLE IN DEVELOPMENT. MAXIMAL
CC ACTIVITY IS OBSERVED AT PH 8.0.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: MAGNESIUM.
CC -1- ENZYME REGULATION: ACTIVATED BY CALCIUM.
CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: 1 (SHOWN HERE), 1A, 1B, 2 AND 3;
CC ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 3 MAY HAVE NO
CC FUNCTIONAL SIGNIFICANCE.
CC -1- TISSUE SPECIFICITY: DETECTED IN RETINA AND RETINAL DERIVED Y-79
CC RETINOBLASTOMA CELLS. ALSO FOUND IN FETAL BRAIN.
CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF023455; AAB82795.1; -
CC EMBL; X97867; CA66461.1; -
CC EMBL; AF027977; AAC05825.1; -
CC EMBL; Z94056; CAB40074.1; -
CC EMBL; AL096700; CAB86407.1; -
CC HSP: P08129; 1PTM.
CC Genew: HSNC:9243; PPEP1.
CC MW: 300109; -
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR000048; IO_region.
CC InterPro: IPR004843; M-peptidase.
```

InterPro: IPR004844; S/T-phosphatase.
 DR Pfam: PF000036; ehand: 3
 DR Pfam: PF00149; Metallophos: 1.
 DR Pfam: PF00612; IQ: 1.
 DR PRINTS: PR00114; STPHPTASE.
 DR Prodom: PD000012; EF-hand: 1.
 DR Prodom: PD000252; S/T-phosphatase: 2.
 DR SMART: SM000054; EPH: 2.
 DR SMART: SM00015; IQ: 1.
 DR SMART: SM00156; PRAC: 1.
 DR PROSITE: PS00018; EF-hand: 2.
 DR PROSITE: PS00125; SER_THR-PHOSPHATASE: 1.
 DR PROSITE: PS50096; IQ: 1.
 KM Hydroxylase; Calcium-binding; Magnesium; Iron; Manganese; Repeat;
 KM Alternative splicing.
 FT DOMAIN 18 43
 FT DOMAIN 121 455
 FT DOMAIN 496 507
 FT CA_BIND 579 590
 FT CA_BIND 619 630
 FT METAL 172 172
 FT METAL 174 174
 FT METAL 201 201
 FT METAL 233 233
 FT ACT_SITE 234 234
 FT METAL 285 285
 FT METAL 403 403
 FT VASPLIC 79 132
 FT VASPLIC 328 355
 FT VASPLIC 356 376
 FT VASPLIC 377 653
 FT VASPLIC 356 417
 FT CONFLICT 367 367
 SQ SEQUENCE 653 AA; 75792 MW; DF7B78C444EE6484 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 653;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 RYEDL 6
 DB 640 RYEDL 644
 RESULT 37
 DNLJ_BACST STANDARD; PRT; 670 AA.
 ID DNLJ_BACST
 AC 087703;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA ligase (EC 6.5.1.2) [Polydeoxyribonucleotide synthase (NAD+)].
 GN LIGA OR LIG.
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.
 OX NCBI_Taxid=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCB 1503;
 RX MEDLINE=99337489; PubMed=10407164;
 RA Brannigan J.A., Ashford S.R., Doherty A.J., Timson D.J., Wigley D.B.;
 RT "Nucleotide sequence, heterologous expression and novel purification
 of DNA ligase from Bacillus stearothermophilus";
 RL Biochim. Biophys. Acta 1432:413-418(1999).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-318.
 RX MEDLINE=99148111; PubMed=10368271;
 RA Singleton M.R., Hakansson K., Timson D.J., Wigley D.B.;
 RT "Structure of the adenylation domain of an NAD+-dependent DNA
 ligase";
 RL Structure 7:35-42(1999).
 CC -!- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER

LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
 CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
 CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
 CC DAMAGED DNA (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (deoxyribonucleotide)(N) +
 CC (deoxyribonucleotide)(N) -> AMP + nicotinamide nucleotide +
 CC (deoxyribonucleotide)(N+M).
 CC -!- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BCT DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ011676; CA09732.1; -
 DR PDB: 1B04; 22-NOV-99.
 DR InterPro: IPR001357; BRCt.
 DR InterPro: IPR004150; DNA_ligase_OB.
 DR InterPro: IPR001679; DNA_ligase.
 DR InterPro: IPR003583; HhH.1.
 DR InterPro: IPR004445; HhH.
 DR InterPro: IPR004149; Znf_DNA_ligase_C4.
 DR Pfam: PF00533; BRCt; 1.
 DR Pfam: PF00633; HhH; 1.
 DR Pfam: PF01653; DNA_ligase_N; 1.
 DR Pfam: PF03119; DNA_ligase_ZBD; 1.
 DR Pfam: PF03120; DNA_ligase_OB; 1.
 DR Prodom: PD003944; DNA_ligase; 1.
 DR SMART: SM00292; BRCt; 1.
 DR SMART: SM00278; HhH; 3.
 DR SMART: SM00532; LIGANC; 1.
 DR TIGRPFAMs: TIGR00575; dnlj; 1.
 DR PROSITE: PS0172; BRCt; 1.
 DR PROSITE: PS01055; DNA_LIGASE_N1; 1.
 DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
 KM Ligase; DNA repair; DNA replication; NAD; 3D-structure.
 FT DOMAIN 589 670
 FT BINDING 114 114
 FT AMP.
 SQ SEQUENCE 670 AA; 74230 MW; B52462314CF9ACF5 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 670;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VYED 5
 DB 122 VYED 126
 RESULT 38
 YDOA_SCHPO STANDARD; PRT; 716 AA.
 ID YDOA_SCHPO
 AC 013730;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C15A10.10 in chromosome I.
 GN SPAC15A10.10.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21648401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mambuli R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Kochet M., Galliard G., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe".
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: 297208; CAB10107.1; -
DR InterPro: IPR001798; Kelch.
DR Pfam: PF01344; Kelch; 2.
KM Hypothetical protein, Repeat.
FT REPEAT 276 322
FT REPEAT 327 381
FT REPEAT 390 439
FT REPEAT 452 499
FT REPEAT 508 558
FT REPEAT 585 635
SQ SEQUENCE 716 AA: 83605 MW: BD2F50D06F5640D CRC64;
Query Match 87.1%; Score 27; DB 1; Length 716;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VYED 5
DB 424 VYED 428

RL Virology 189:304-316(1992).
CC -1- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC HSV-1 56, EBV BRFL1, HCMV UL104, AND VAV 54.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M86664; AB02491.1; -
DR PIR: B36801; WZBE8.
DR InterPro: IPR002660; Herpes_UL6.
DR Pfam: PF01763; Herpes_UL6; 1.
DR Prodom: PD003210; Herpes_UL6; 1.
SQ SEQUENCE 753 AA: 83992 MW: C5E118F78BBD203 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 753;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 108 RYEDL 112

RESULT 40
PARC_CAUCR STANDARD; PRT: 759 AA.
ID 054478:
AC 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Topoisomerase IV subunit A (EC 5.99.1.1).
GN PARC OR C01566.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group.
CC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=98086097; PubMed=9426128;
RA Ward D.V., Newton A.;
RT "Requirement of topoisomerase IV *parc* and *pare* genes for cell cycle
RT progression and developmental regulation in *Caulobacter crescentus*."
RL Mol. Microbiol. 26:897-910(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RA Ward D.V., Newton A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Iaub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Pirocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Hall D.E.,
RA Kolonay J.F., Salt J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uiterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of *Caulobacter crescentus*."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: TOPOISOMERASE IV IS ESSENTIAL FOR CHROMOSOME
CC SEGREGATION. IT HAS RELAXATION OF SUPERCOILED DNA ACTIVITY.
CC PERFORMS THE DECATENATION EVENTS REQUIRED DURING THE REPLICATION
CC OF A CIRCULAR DNA MOLECULE (BY SIMILARITY).

```

CC -1- SUBUNIT: COMPOSED OF TWO SUBUNITS: PARC AND PARF.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: STRONG, WITH THE A SUBUNIT OF GINASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94696; AAC38043.1; -.
DR EMBL: U86302; AAF14339.1; -.
DR EMBL: AF005831; AAK23545.1; -.
DR HSP: P09097; IAB4.
DR TIGR: CC1566; -.
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV.1.
DR ProDom: PD000742; DNA_topoisomIV.1.
DR SMART: SM00434; TOP4C; 1.
DR TIGRfams: TIGR01062; parC_Gneg; 1.
DR Topoisomerase: isomerase; DNA-binding; Complete proteome.
FT ACT_SITE 132 132 DNA_CLEAVAGE (BY SIMILARITY).
FT CONFLICT 498 498 R -> A (IN REF. 1 AND 2).
SQ SEQUENCE 759 AA; 83521 MW; 65DDA34154A347D4 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 759;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
DB 408 VRYED 412

RESULT 41
DIVL CAUCR STANDARD; PRT; 769 AA.
ID DIVL CAUCR
AC Q9RQ09; Q9A252;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensor protein divL (EC 2.7.3.-).
GN DIVL OR CC3484.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter.
CC NCBI_Taxid=155892;
RN [1]
RP SEQUENCE FROM N.A. AND AUTOPHOSPHORYLATION SITE.
RC STRAIN=ATCC 19089 / CB15;
RA MEDLINE=20027501; PubMed=10557274;
RA Wu J., Ohta N., Zhao J.L., Newton A.;
RA "A novel bacterial tyrosine kinase essential for cell division and
RA differentiation.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:13068-13073(1999).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647;
RA Niernan W.C., Feldblum T.V., Laub M.R.K., Ohta N., Maddock J.R.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.D., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utechtack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus.";
RT Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Required for cell division and growth. It catalyzes the
CC phosphorylation of CtrA and activates transcription in vitro of
CC the cell cycle-regulated flhF promoter.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: Autophosphorylated.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF083422; AAF08344.2; -.
DR EMBL: AF006007; AAK25446.1; -.
DR TIGR: CC3484; -.
DR ProSite: Q9RQ09; -.
DR InterPro: IPR004358; Bact_sens-PR-C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR004359; His_kinA_sig.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; signal; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00388; HSKA; 1.
DR TIGRfams: TIGR00229; sensory_box; 2.
DR PROSITE: PS50109; His_kin; 1.
DR Sensory transduction: Transferase; Kinase; Transmembrane;
KW Phosphorylation; Complete proteome.
FT TRANSMEM 5 26 POTENTIAL.
FT DOMAIN 547 758 HISTIDINE KINASE.
FT DOMAIN 9 221 ALA-RICH.
FT MOD_RES 550 550 PHOSPHORYLATION (AUTO-).
FT CONFLICT 200 200 Q -> H (IN REF. 1).
FT CONFLICT 216 216 V -> E (IN REF. 1).
SQ SEQUENCE 769 AA; 82736 MW; 002B24248F18A57EF CRC64;

Query Match 87.1%; Score 27; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 351 RYEDL 355

RESULT 42
UL06 VZVD STANDARD; PRT; 769 AA.
ID UL06 VZVD
AC P09302;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Varion gene 54 protein.
GN 54.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
CC NCBI_Taxid=10338;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=8630657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RA "The complete DNA sequence of varicella-zoster virus.";
RT J. Gen. Virol. 67:1759-1816(1986)
CC -1- FUNCTION: PRESUMED VIRION PROTEIN. POSSIBLE ROLE IN DNA
CC PACKAGING (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC EBV-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04370; CAA27937.1; -
DR PIR; B27215; W2BE54.
DR InterPro; IPR002660; Herpes_UL6.
DR Pfam; PF01763; Herpes_UL6; 1.
DR Prodom; PD003210; Herpes_UL6; 1.
SQ SEQUENCE 769 AA; 86780 MM; 5ABD7EDA6D783BCF CRC64;

Query Match 87.1%; Score 27; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 43
ORP5_MOUSE STANDARD; PRT; 874 AA.
AC Q9ER64; Q99NE5; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxytelin binding protein-related protein 5 (OSBP-related protein 5)
DE (Orp-5) (Oxytelin-binding protein homologue 1).
GN OSBP5 OR OSBP2 OR OSBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=129/SV.
RC MEDLINE=20519229; Pubmed=11063728;
RA Engemann S., Stroedlecke M., Paulsen M., Franck O., Reinhardt R.,
RA Lane N., Reik W., Walter J.;
RT "Sequence and functional comparison in the Beckwith-Wiedemann region:
RT implications for a novel imprinting centre and extended imprinting.";
RL Hum. Mol. Genet. 9:2691-2706(2000).
CC -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ278263; CAC16404.2; ALT_INIT.
DR EMBL; AJ276505; CAC27351.1; ALT_INIT.
DR MGD; MGI:1930265; Osbp2.
DR InterPro; IPR000648; Oxytelin_BP.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01237; Oxytelin_BP; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS01013; OSBP; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Lipid transport; transport.
FT DOMAIN 126 243 PH.
FT CONFLICT 37 44 ENELGPIT -> MSLVPSQ (IN REF. 1;
FT CAC27351).
SQ SEQUENCE 874 AA; 98921 MW; FBC41FA8E219F5E3 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 874;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 707 RYEDL 711

RESULT 44
PKC1_ASPNG STANDARD; PRT; 1096 AA.
AC Q00078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase C-like (EC 2.7.1.1).
GN PKCA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96158841; Pubmed=8569684;
RA Morawetz R., Lendenfeld T., Mischak H., Muehlbauer M., Gruber F.,
RA Goodnight J., de Graaff L.H., Visser J., Mushinski J.F.,
RA Kubicek C.P.;
RT "Cloning and characterisation of genes (pkc1 and pkca) encoding
RT protein kinase C homologues from Trichoderma reesei and Aspergillus
RT niger.";
RL Mol. Genet. 250:17-28(1996).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHOSPHO-ESTER AND DAG
CC PKC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10549; AA97433.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000861; REM_repeat.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00130; DAG_PE-bind; 2.
DR Pfam; PF00433; Pkinase_C; 1.
DR Prodom; PR000001; Euk_pkinase; 1.
DR SMART; SM00109; C1; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00074; HRI; 2.
DR SMART; SM00133; S_TR_X; 1.
DR SMART; SM00220; S_TR_X; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE; PS50061; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS01017; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS01018; PROTEIN_KINASE_ST; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding; zinc;
KW Phospho-ester binding; Repeat.
FT DOMAIN 460 507 PHOSPHO-ESTER AND DAG BINDING 1;
FT DOMAIN 528 577 PHOSPHO-ESTER AND DAG BINDING 2;
FT DOMAIN 771 1030 PROTEIN KINASE.
FT NP_BIND 777 785 ATP (BY SIMILARITY).
FT BINDING 800 800 ATP (BY SIMILARITY).

FT ACU_SITE 896 896 BY SIMILARITY
SO SEQUENCE 1096 AA; 122234 MW; 859B2B3530D5B08 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 1096;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 218 RYEDL 222

RESULT 45
RPOB_THEAO STANDARD; PRT: 1119 AA.
ID RPOB_THEAO
AC Q9KWT7; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
beta chain) (RNA polymerase beta subunit).
GN RPOB.
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=271;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE=99428144; PubMed=10499798;
RA Zhang G., Campbell E.A., Minkhin L., Richter C., Severinov K.,
RA Darst S.A.;
RT "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A
resolution."
RL Cell 98:811-824(1999).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
(RNA)(N).
CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
BETA' CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL, Y19223; CAB65465.2; -.
CC PDB; 1HQW; 07-FEB-01.
CC InterPro; IPR001572; RNA_POL_B.
CC Pfam; PF00562; RNA_POL_B; 2.
CC PROSITE; PS01166; RNA_POL_BETA; 1.
CC Transferrase; Transcription; DNA-directed RNA polymerase;
KW 3D-structure 1119 AA; 124757 MW; 2CF66FA79C77F33 CRC64;
SQ SEQUENCE 1119 AA; 124757 MW; 2CF66FA79C77F33 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 1119;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRYED 5
DB 613 VRYED 617

RESULT 46
KPC1_COCHC STANDARD; PRT: 1174 AA.
ID KPC1_COCHC

AC 042632;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase C-like (EC 2.7.1.1).
GN KPC1.
OS Cochliobolus heterostrophus (Drechslera maydis).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; Cochliobolus.
OX NCBI_TaxID=5016;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 48329 / C2;
RA Oeser B.M., Yoder O.C.;
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; Y15839; CAA75801.1; -.
CC HSSP; 063450; 1A06.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002219; DAG_PE_BIND.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000861; REM_repeat.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00130; DAG-pe-bind; 2.
DR Pfam; PF00433; pkinase_C; 1.
DR Pfam; PF02185; HRI; 2.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00109; C1; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00074; HRI; 2.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK_X; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Zinc;
KW Phorbol-ester binding; Repeat.
FT DOMAIN 459 506
FT DOMAIN 527 576
FT DOMAIN 704 709
FT DOMAIN 785 792
FT DOMAIN 849 1108
FT NP_BIND 855 863
FT BINDING 878 878
FT ACT_SITE 974
SQ SEQUENCE 1174 AA; 130506 MW; 26A4ADDA2849F37C CRC64;
Query Match 87.1%; Score 27; DB 1; Length 1174;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 219 RYEDL 223

RESULT 47

PROTEIN ID	STANDARD	PRT:	1393 AA.
PROT_CHLPPN			
AC	092999: Q9JUS7, Q9K211;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase beta' chain) (RNA polymerase beta' subunit).		
GN	RPOC OR CPN0082 OR CP0593		
OS	Chlamydia pneumoniae (Chlamydochlamydia pneumoniae).		
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydochlamydia.		
CX	NCBI_TaxID=83558;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CML029;		
RA	MEDLINE=99206606; PubMed=10192388;		
RA	Altman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., Stephens R.S.;		
RT	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";		
RL	Nat. genet. 21:385-389(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-AR39;		
RA	Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.R., Peterson J., Ulterback T., Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;		
RT	"Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.";		
RL	Nucleic Acids Res. 28:1397-1406(2000).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-J138;		
RA	MEDLINE=20330349; PubMed=10871362;		
RA	Shira M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;		
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CML029 from USA.";		
RL	Nucleic Acids Res. 28:2311-2314(2000).		
CC	-1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.		
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).		
CC	-1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA CHAIN.		
CC	-1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AE001593; AAD18235.1; -		
DR	EMBL; AE002228; AAF38501.1; AUT_INIT.		
DR	EMBL; AP002545; BAA98292.1; -		
DR	HSSP; O9K063; IHOM.		
DR	TIGR; CP0693;		
DR	InterPro; IPR000722; RNA_POL_A.		
DR	InterPro; IPR002879; RNA_POL_A2.		
DR	Pfam; PF00623; RNA_POL_A; 1.		
DR	Pfam; PF01854; RNA_POL_A2; 2.		
KW	Transferrase; DNA-directed RNA polymerase; Transcription;		
KW	Complete proteome.		
FT	CONFLICT 1031 A->G (IN REF. 1).		
SD	SEQUENCE 1393 AA; 154900 MW; E0734EE236C66FE8 CRC64;		

[illegible]

RESULT 50
DYH9_HUMAN STANDARD; PRT; 4486 AA.
ID DYH9_HUMAN STANDARD; PRT; 4486 AA.
AC Q9NVC9; Q9NVC9; 095494; (Created)
DI 16-OCT-2001 (Rel. 40, last sequence update)
DI 16-OCT-2001 (Rel. 40, last sequence update)
DI 16-OCT-2001 (Rel. 40, last sequence update)
DE Ciliary dynein heavy chain (Axonemal dynein heavy chain) (Dynein heavy chain 9) OR DNAL1.
GN DNAL1 OR DNAL1L OR DNAL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Reed W., Moats-Staats B.M., Carson J.L., Leigh M.W., Collier A.M.;
RT "A ciliary dynein heavy chain whose expression is upregulated in differentiating airway epithelium."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Nasal epithelium;
RA Bartoloni L., Blouin J.L., Maiti A., Sainsbury A., Rossler C., Gehrig C., She J.X., Marron M.P., Lander E., Meeks M., Chung E., Jorissen M., Scott H.S., Delozier Blanchet C.D., Gardiner M., Antonarakis S.E.;
RT "Axonemal beta heavy chain dynein DNAL1: cDNA sequence, genomic structure and investigation of its role in primary ciliary dyskinesia."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1874-1974 FROM N.A.
RC TISSUE=Nasal polyps;
RA Maiti A.K., Maiti M.G., Jorissen M., Volz A., Ziegler A., Bouvagnet P.;
RT "Chromosomal localization of human dynein heavy chain genes."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RA MEDLINE=20558134; PubMed=1104725;
REED W., Carson J.L., Moats-Staats B.M., Lucier T., Hu P.C., Brighton L., Gambling T.M., Huang C.H., Leigh M.W., Collier A.M.;
RT "Characterization of an axonemal dynein heavy chain expressed early in airway epithelial ciliogenesis."
RL Am. J. Respir. Cell Mol. Biol. 23:734-741(2000).
CC -1- FUNCTION: FORCE GENERATING PROTEIN OF RESPIRATORY CILIA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE ACTIVITY.
CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF INTERMEDIATE AND LIGHT CHAINS.
CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF257737; AAF69004.1; -;
CC EMBL: AJ404468; CAB94756.1; -;
CC EMBL: AJ132088; CA10561.1; -;
CC Gene: HGNC:2953; DNAL1.
CC MIM: 603330; -;
CC InterPro: IPR004273; Dynein_heavy.
CC Pfam: PF03026; Dynein_heavy; 1.
CC Motor protein: Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 381 410 COILED COIL (POTENTIAL).
FT DOMAIN 504 529 COILED COIL (POTENTIAL).

FT	DOMAIN	639	662	COILED COIL (POTENTIAL).
FT <td>DOMAIN</td> <td>752</td> <td>823</td> <td>COILED COIL (POTENTIAL).</td>	DOMAIN	752	823	COILED COIL (POTENTIAL).
FT <td>DOMAIN</td> <td>1326</td> <td>1355</td> <td>COILED COIL (POTENTIAL).</td>	DOMAIN	1326	1355	COILED COIL (POTENTIAL).
FT <td>DOMAIN</td> <td>3051</td> <td>3154</td> <td>COILED COIL (POTENTIAL).</td>	DOMAIN	3051	3154	COILED COIL (POTENTIAL).
FT <td>DOMAIN</td> <td>3285</td> <td>3341</td> <td>COILED COIL (POTENTIAL).</td>	DOMAIN	3285	3341	COILED COIL (POTENTIAL).
FT <td>DOMAIN</td> <td>3640</td> <td>3675</td> <td>COILED COIL (POTENTIAL).</td>	DOMAIN	3640	3675	COILED COIL (POTENTIAL).
FT <td>NP_BIND</td> <td>1870</td> <td>1877</td> <td>ATP (POTENTIAL).</td>	NP_BIND	1870	1877	ATP (POTENTIAL).
FT <td>NP_BIND</td> <td>2151</td> <td>2158</td> <td>ATP (POTENTIAL).</td>	NP_BIND	2151	2158	ATP (POTENTIAL).
FT <td>NP_BIND</td> <td>2478</td> <td>2485</td> <td>ATP (POTENTIAL).</td>	NP_BIND	2478	2485	ATP (POTENTIAL).
FT <td>NP_BIND</td> <td>2825</td> <td>2832</td> <td>ATP (POTENTIAL).</td>	NP_BIND	2825	2832	ATP (POTENTIAL).
FT <td>CONFLICT</td> <td>2505</td> <td>2505</td> <td>L -> V (IN REF. 2).</td>	CONFLICT	2505	2505	L -> V (IN REF. 2).
FT <td>CONFLICT</td> <td>3678</td> <td>3678</td> <td>T -> A (IN REF. 2).</td>	CONFLICT	3678	3678	T -> A (IN REF. 2).
FT <td>CONFLICT</td> <td>4374</td> <td>4374</td> <td>I -> M (IN REF. 2).</td>	CONFLICT	4374	4374	I -> M (IN REF. 2).
SQ <td>SEQUENCE</td> <td>4486 AA;</td> <td>511927 MW;</td> <td>996EDFEDEB0B3EB1 CRC64;</td>	SEQUENCE	4486 AA;	511927 MW;	996EDFEDEB0B3EB1 CRC64;

Query Match
Best Local Similarity 57.1%; Score 27; DB 1; Length 4486;
Matches 5; Conservative 100.0%; Pred. No. 1.6e+03; Indels 0; Gaps 0;
DB 258 RYEDL 262

Search completed: February 20, 2003, 13:32:38
Job time: 16.5714 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: February 20, 2003, 13:29:11 ; Search time 8.57143 Seconds

(without alignments)
144.233 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31
Sequence: 1 VAYEDL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database :
1: SP:archae:21:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mnc:*
8: SP:organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP:tvirus:*
16: SP:bacteriopl:*
17: SP:archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	114 11 Q9D0K5	Q9D0K5 mus musculus
2	31	100.0	183 5 Q9VMC2	Q9VMC2 drosophila
3	31	100.0	281 17 Q980M6	Q980M6 sulfolobus
4	31	100.0	331 12 Q11316	Q11316 molluscum c
5	31	100.0	386 4 Q9Y5R3	Q9Y5R3 mus sapien
6	31	100.0	395 11 Q9Q0P4	Q9Q0P4 mus musculu
7	31	100.0	411 4 Q43916	Q43916 homo sapien
8	31	100.0	411 11 Q9E0C0	Q9E0C0 mus musculu
9	31	100.0	428 16 Q9RVQ3	Q9RVQ3 mus musculu
10	31	100.0	484 4 Q90ED5	Q90ED5 delinococcus
11	31	100.0	486 5 Q95TN9	Q95TN9 homo sapien
12	31	100.0	530 4 Q9Y6F2	Q9Y6F2 drosophila
13	31	100.0	530 11 Q88276	Q88276 mus musculu
14	31	100.0	531 4 Q9Y4C5	Q9Y4C5 homo sapien
15	31	100.0	537 5 Q9Y158	Q9Y158 drosophila
16	31	100.0	631 16 Q25212	Q25212 helicobacte

17	31	100.0	649 5 Q9V884	Q9V884 drosophila
18	31	100.0	748 12 Q98200	Q98200 molluscum c
19	31	100.0	973 5 Q45412	Q45412 caenorhabdi
20	31	100.0	6260 2 Q54299	Q54299 streptomyce
21	30	96.8	194 17 Q82TB3	Q82TB3 pyrobaculum
22	30	96.8	199 11 Q9D806	Q9D806 mus musculu
23	30	96.8	363 5 Q9VMC4	Q9VMC4 drosophila
24	30	96.8	540 17 Q8T0C3	Q8T0C3 methanosarc
25	30	96.8	579 16 Q9A7T6	Q9A7T6 caulobacter
26	30	96.8	595 5 Q8SRF6	Q8SRF6 encephalito
27	30	96.8	1322 10 Q23559	Q23559 arbidopsis
28	29	93.5	234 16 Q9A994	Q9A994 caulobacter
29	29	93.5	247 12 Q9JF80	Q9JF80 ectromelia
30	29	93.5	303 16 Q93UE6	Q93UE6 streptomyce
31	29	93.5	355 16 Q926E4	Q926E4 rhizobium m
32	29	93.5	393 10 Q81244	Q81244 cucumis mel
33	29	93.5	405 10 Q22818	Q22818 arbidopsis
34	28	90.3	154 2 P96458	P96458 streptomyce
35	28	90.3	195 17 Q28741	Q28741 archaeoglob
36	28	90.3	228 17 Q976N0	Q976N0 sulfolobus
37	28	90.3	233 16 Q9CEP1	Q9CEP1 lactococcus
38	28	90.3	273 2 Q9RC18	Q9RC18 bacillus sp
39	28	90.3	273 12 Q9YRQ1	Q9YRQ1 atelline her
40	28	90.3	278 2 Q9ET84	Q9ET84 bacillus am
41	28	90.3	306 16 Q91732	Q91732 pseudomonas
42	28	90.3	333 10 Q82330	Q82330 arbidopsis
43	28	90.3	346 5 Q45982	Q45982 caenorhabdi
44	28	90.3	354 10 Q9FL45	Q9FL45 arbidopsis
45	28	90.3	359 10 Q940P4	Q940P4 arbidopsis
46	28	90.3	388 11 Q9R111	Q9R111 mus musculu
47	28	90.3	388 11 Q9WUE5	Q9WUE5 mus musculu
48	28	90.3	391 17 Q97B92	Q97B92 thermoplas
49	28	90.3	458 17 Q9V129	Q9V129 pyrococcus
50	28	90.3	472 2 Q93N59	Q93N59 coxiella bu
51	28	90.3	472 11 Q81819	Q81819 mus musculu
52	28	90.3	474 11 Q9Q2L2	Q9Q2L2 ratus norv
53	28	90.3	479 4 Q75099	Q75099 homo sapien
54	28	90.3	484 11 Q99NB0	Q99NB0 mus musculu
55	28	90.3	484 11 Q9EP78	Q9EP78 mus musculu
56	28	90.3	486 4 Q9NS84	Q9NS84 homo sapien
57	28	90.3	486 4 Q75667	Q75667 homo sapien
58	28	90.3	503 17 Q980L7	Q980L7 sulfolobus
59	28	90.3	507 16 Q99WJ7	Q99WJ7 staphylococ
60	28	90.3	541 16 Q98308	Q98308 rhizobium l
61	28	90.3	612 10 Q9SD68	Q9SD68 arbidopsis
62	28	90.3	622 17 Q93736	Q93736 pyrococcus
63	28	90.3	672 16 Q9P8S8	Q9P8S8 streptomyce
64	28	90.3	697 13 Q90YX2	Q90YX2 discoglossu
65	28	90.3	718 12 Q91BWB	Q91BWB turkey herp
66	28	90.3	722 12 Q9E6R0	Q9E6R0 turkey herp
67	28	90.3	788 12 Q9QCD7	Q9QCD7 marek's dis
68	28	90.3	793 5 Q9VBX0	Q9VBX0 drosophila
69	28	90.3	828 5 Q8SX93	Q8SX93 drosophila
70	28	90.3	875 5 Q9N396	Q9N396 caenorhabdi
71	28	90.3	943 5 Q9VFE3	Q9VFE3 drosophila
72	28	90.3	1078 17 Q8TPH6	Q8TPH6 methanosarc
73	28	90.3	1107 16 Q8XDP6	Q8XDP6 escherichia
74	28	90.3	1658 10 Q8S622	Q8S622 cryza sativ
75	28	90.3	2258 12 Q89238	Q89238 hyman parat
76	27	87.1	52 2 Q9RGS3	Q9RGS3 staphylococ
77	27	87.1	73 16 Q9KAO6	Q9KAO6 bacillus ha
78	27	87.1	108 17 Q9YDE8	Q9YDE8 aeropyrum p
79	27	87.1	112 2 Q9L4E2	Q9L4E2 uncultured
80	27	87.1	112 2 Q9L4E1	Q9L4E1 uncultured
81	27	87.1	127 3 Q9L479	Q9L479 uncultured
82	27	87.1	127 3 Q9P895	Q9P895 emericella
83	27	87.1	128 5 Q9U683	Q9U683 culic pipie
84	27	87.1	133 17 Q28799	Q28799 archaeoglob
85	27	87.1	149 5 Q9U019	Q9U019 plasmodium
86	27	87.1	157 2 Q9L419	Q9L419 corynebacte
87	27	87.1	163 2 Q9EUT9	Q9EUT9 klebsiella
88	27	87.1	163 2 Q9EUT8	Q9EUT8 klebsiella
89	27	87.1	163 2 Q9EXR7	Q9EXR7 klebsiella

90	27	87.1	163	2	Q9EXT6	Q9ext6 klebsiella	163	27	87.1	504	16	Q92LD7	Q92ld7 rhizobium m
91	27	87.1	163	2	Q9EXT5	Q9ext5 klebsiella	163	27	87.1	516	2	Q9AN11	Q9an11 bradyrhizob
92	27	87.1	163	2	Q9EXT4	Q9ext4 klebsiella	163	27	87.1	516	2	Q93HS0	Q93hs0 bradyrhizob
93	27	87.1	163	2	Q9EXT3	Q9ext3 klebsiella	166	27	87.1	519	2	Q93N44	Q93n44 coxiella bu
94	27	87.1	181	17	Q8SZ0	Q8sz0 pyrobaculum	167	27	87.1	520	5	Q45411	Q45411 caenorhabdi
95	27	87.1	188	4	Q96OB9	Q96ob9 homo sapien	168	27	87.1	522	2	Q9KHK1	Q9kx1 streptomyces
96	27	87.1	189	17	Q96Y3	Q96y3 sulfobus	169	27	87.1	525	16	Q9HUA5	Q9hu5 pseudomonas
97	27	87.1	200	3	Q9HDG9	Q9hdg9 cryptococcu	170	27	87.1	527	17	Q29232	Q29232 archaeoglob
98	27	87.1	200	3	Q9HGS6	Q9hgs6 cryptococcu	171	27	87.1	543	2	Q9K2F3	Q9k2f3 uncultured
99	27	87.1	200	3	Q9HGS5	Q9hgs5 cryptococcu	172	27	87.1	550	3	Q9P8P1	Q9p8p1 cryptococcu
100	27	87.1	200	3	Q9HGS4	Q9hgs4 cryptococcu	173	27	87.1	550	4	Q9NTN2	Q9ntn2 homo sapien
101	27	87.1	200	16	Q922M8	Q922m8 rhizobium m	174	27	87.1	551	4	Q95678	Q95678 homo sapien
102	27	87.1	203	16	Q9ZHC2	Q9zhc2 streptococc	175	27	87.1	551	4	Q9NSA9	Q9nsa9 homo sapien
103	27	87.1	203	16	Q97PF7	Q97pf7 streptococc	176	27	87.1	554	3	Q9P399	Q9p399 neurospora
104	27	87.1	225	16	Q66072	Q66072 canine herp	177	27	87.1	554	5	Q8SR05	Q8sr05 encephalito
105	27	87.1	225	16	Q66755	Q66755 aquifex aeo	178	27	87.1	557	10	Q9XIC5	Q9xics arabidopsis
106	27	87.1	236	10	Q65802	Q65802 ceratopter	179	27	87.1	561	16	Q8U752	Q8u752 agrobacteri
107	27	87.1	239	10	Q8TQ65	Q8tq65 methanosarc	180	27	87.1	584	10	Q9AVZ7	Q9avz7 cryza sativ
108	27	87.1	241	10	Q9FKD6	Q9fk65 arabidopsis	181	27	87.1	592	3	Q59900	Q59900 cryptococcu
109	27	87.1	245	2	Q51405	Q51405 pseudomonas	182	27	87.1	592	16	Q98H21	Q98hz1 rhizobium 1
110	27	87.1	253	7	P79522	P79522 homo sapien	183	27	87.1	604	4	Q9NQ27	Q9nq27 homo sapien
111	27	87.1	255	16	Q99R22	Q99r22 staphylococ	184	27	87.1	605	4	Q8TB00	Q8tbu0 homo sapien
112	27	87.1	275	16	Q8ZMR1	Q8zmr1 salmonella	185	27	87.1	605	16	Q8XAK0	Q8xak0 escherichia
113	27	87.1	279	2	Q51406	Q51406 pseudomonas	186	27	87.1	606	11	Q9ETJ0	Q9etj0 mus musculu
114	27	87.1	283	16	Q9JUK4	Q9juk4 neisseria m	187	27	87.1	617	16	Q92206	Q92206 bacillus su
115	27	87.1	284	12	Q66105	Q66105 citrus leaf	188	27	87.1	624	17	Q8U410	Q8u410 pyrococcus
116	27	87.1	298	10	Q9LV42	Q9lv42 arabidopsis	189	27	87.1	630	16	Q8UER9	Q8uer9 agrobacteri
117	27	87.1	307	16	Q92VG4	Q92vg4 rhizobium m	190	27	87.1	631	16	Q9ZM07	Q9zm07 helicobacte
118	27	87.1	308	4	Q9BMA0	Q9bma0 homo sapien	191	27	87.1	653	12	Q64906	Q64906 acetalphne
119	27	87.1	312	17	Q9YD31	Q9y31 aeropyrum p	192	27	87.1	666	16	Q8Y828	Q8y828 listeria mo
120	27	87.1	315	5	Q9N3F7	Q9n3f7 caenorhabdi	193	27	87.1	685	16	Q92TJ0	Q92tj0 rhizobium m
121	27	87.1	324	16	Q98LQ4	Q98lq4 rhizobium 1	194	27	87.1	693	2	Q91822	Q91822 enterococcu
122	27	87.1	325	5	Q9VN98	Q9vn98 drosophila	195	27	87.1	693	4	Q9UPU1	Q9upu1 homo sapien
123	27	87.1	334	16	Q9K689	Q9k689 bacillus ha	196	27	87.1	694	16	Q926R0	Q926r0 chlamydia p
124	27	87.1	338	12	Q9MWZ1	Q9wmz1 dioscorea a	197	27	87.1	706	5	Q9Y417	Q9y417 drosophila
125	27	87.1	351	2	Q52572	Q52572 amycolatops	198	27	87.1	713	10	Q9FT72	Q9ft72 arabidopsis
126	27	87.1	352	16	Q9K9N7	Q9k9n7 bacillus ha	199	27	87.1	747	12	Q93296	Q93296 equine herp
127	27	87.1	355	2	Q51154	Q51154 neisseria m	200	27	87.1	769	2	Q9R0Q9	Q9r0q9 caulobacter
128	27	87.1	359	2	Q9F016	Q9f016 xanthomonas	201	27	87.1	773	16	Q9A252	Q9a252 caulobacter
129	27	87.1	369	2	Q9L307	Q9l307 streptomyces	202	27	87.1	785	17	Q9YCC8	Q9yc8 aeropyrum p
130	27	87.1	378	12	Q9R8Z5	Q9r8z5 pseudomonas	203	27	87.1	809	16	Q9A704	Q9a7q4 caulobacter
131	27	87.1	378	12	Q91EU9	Q91eu9 cydia pomon	204	27	87.1	850	4	Q9ULE7	Q9ule7 homo sapien
132	27	87.1	379	2	Q68541	Q68541 cellvibrio	205	27	87.1	874	11	Q8R510	Q8r510 mus musculu
133	27	87.1	380	13	Q90XHO	Q90xho brachydanio	206	27	87.1	882	16	Q92JW2	Q92jw2 rhizobium m
134	27	87.1	381	11	Q99WH7	Q99wh7 mus musculu	207	27	87.1	882	16	Q92JW2	Q92jw2 rhizobium m
135	27	87.1	389	16	Q9K464	Q9k464 streptomyces	208	27	87.1	891	3	Q9US57	Q9us57 schizosacch
136	27	87.1	393	13	Q9PTM7	Q9ptm7 brachydanio	209	27	87.1	926	10	Q81805	Q81805 arabidopsis
137	27	87.1	393	13	Q9DDE0	Q9d93 homo sapien	210	27	87.1	976	16	Q9ANV4	Q9anv4 caulobacter
138	27	87.1	395	4	Q9GZ3	Q9gzx3 homo sapien	211	27	87.1	977	16	Q9BNV6	Q9bnv6 rhizobium 1
139	27	87.1	399	1	Q9DBO5	Q9db95 mus musculu	212	27	87.1	991	3	Q96VF6	Q96vf6 tuber magna
140	27	87.1	411	4	Q9GZS9	Q9gz95 homo sapien	213	27	87.1	1004	3	Q9HE23	Q9he23 neurospora
141	27	87.1	418	17	Q8TVF2	Q8tvf2 methanopyru	214	27	87.1	1079	5	Q17566	Q17566 caenorhabdi
142	27	87.1	424	16	Q9KTY6	Q9ky6 streptomyces	215	27	87.1	1094	5	Q23915	Q23915 dictyosteli
143	27	87.1	429	16	Q9A018	Q9a018 streptococc	216	27	87.1	1108	13	Q9PMD0	Q9pmd0 tetradon f
144	27	87.1	439	2	Q8RME3	Q8rme3 streptococc	217	27	87.1	1119	2	Q8ROE9	Q8roe9 thermus the
145	27	87.1	439	16	Q8XD63	Q8xd63 escherichia	218	27	87.1	1136	3	Q9HGK8	Q9hgk8 tuber borch
146	27	87.1	445	17	Q9HK59	Q9hk59 thermoplasm	219	27	87.1	1137	5	Q93250	Q93250 caenorhabdi
147	27	87.1	447	4	Q96AK9	Q96ak9 homo sapien	220	27	87.1	1157	3	Q9HE10	Q9he10 blumeria gr
148	27	87.1	453	16	Q93MB2	Q93mb2 clostridium	221	27	87.1	1265	4	Q9E837	Q9e837 homo sapien
149	27	87.1	455	17	Q8U0L4	Q8u0l4 pyrococcus	222	27	87.1	1352	17	Q58837	Q58837 pyrococcus
150	27	87.1	456	16	Q9F9K9	Q9f9k9 pisciricket	223	27	87.1	1359	10	Q9SMW4	Q9smw4 arabidopsis
151	27	87.1	462	16	Q8UTK8	Q8utk8 agrobacteri	224	27	87.1	1405	10	Q9LH88	Q9lh88 arabidopsis
152	27	87.1	472	2	Q43941	Q43941 acinetobact	225	27	87.1	1678	13	Q8UWH8	Q8uwh8 cryzias lat
153	27	87.1	474	16	Q9KSM2	Q9ksm2 vibrio chol	226	27	87.1	1833	5	Q9VM67	Q9vm67 drosophila
154	27	87.1	476	16	Q9KRN0	Q9krn0 vibrio chol	227	27	87.1	2173	5	Q9U622	Q9u622 drosophila
155	27	87.1	478	5	Q8SAX1	Q8sax1 drosophila	228	27	87.1	2257	5	Q9V132	Q9v132 drosophila
156	27	87.1	488	2	Q87253	Q87253 lactococcus	229	27	87.1	3187	5	Q9BLV5	Q9blv5 leishmania
157	27	87.1	488	2	Q48617	Q48617 lactococcus	230	27	87.1	5171	4	Q8WVK9	Q8wvk9 homo sapien
158	27	87.1	488	16	Q9CHY0	Q9chy0 lactococcus	231	27	87.1	10223	2	Q54296	Q54296 streptomyces
159	27	87.1	499	16	Q44318	Q44318 anabaena sp	232	27	87.1	105	17	Q97U37	Q97u37 sulfobus
160	27	87.1	500	16	Q8YKN8	Q8ykn8 anabaena sp	233	27	87.1	117	2	Q9L4C9	Q9l4c9 gamma prote
161	27	87.1	500	16	Q9ANB5	Q9anb5 bradyrhizob	234	27	87.1	117	2	Q9L4C9	Q9l4c9 gamma prote
162	27	87.1	504	16	Q9ABF5	Q9abf5 caulobacter	235	27	87.1	121	2	P95818	P95818 streptococc

236	26	83.9	123	6	Q9GMX9	Q9gmx9 macaca fasc	309	26	83.9	662	5	062239	062299 caenorhabdi
237	26	83.9	123	9	Q9G006	Q9g006 bacterioplasm	310	26	83.9	667	2	08VUH1	08vuh1 streptococc
238	26	83.9	133	17	Q97AV4	Q97av4 thermoplasm	311	26	83.9	671	16	09KVH9	09kvh9 vibrio chol
239	26	83.9	136	11	Q9PIU9	Q9piu9 campylobact	312	26	83.9	686	10	09C602	09c602 arabidopsis
240	26	83.9	137	16	089076	089076 mus musculu	313	26	83.9	691	10	08VXZ2	08vxz2 arabidopsis
241	26	83.9	141	4	Q8WURO	Q8wuro mus musculu	314	26	83.9	695	5	096441	096441 dirosophila
242	26	83.9	141	11	Q9DAX5	Q9dax5 homo sapien	315	26	83.9	695	5	09U8D7	09u8d7 chymomyza a
243	26	83.9	147	17	Q82WP9	Q82wp9 mus musculu	316	26	83.9	699	17	08TWT4	08tw4 methanopyru
244	26	83.9	148	10	Q8VZ50	Q8vz50 pyrobaculum	317	26	83.9	700	5	08WS94	08ws94 caenorhabdi
245	26	83.9	148	10	Q94Z44	Q94z44 arabidopsis	318	26	83.9	702	17	09HMZ6	09hmz6 thermoplasm
246	26	83.9	149	6	Q9BGT5	Q9bgt5 arabidopsis	319	26	83.9	706	5	Q9NDP3	Q9ndp3 andes dictyosteli
247	26	83.9	157	6	Q9K408	Q9k408 macaca fasc	320	26	83.9	715	10	Q8S2D3	Q8s2d3 oryza sativ
248	26	83.9	158	15	Q41878	Q41878 human immun	321	26	83.9	721	16	Q9A4D9	Q9a4d9 caulobacter
249	26	83.9	176	2	Q9K4W9	Q9k4w9 lactobacilli	322	26	83.9	727	4	08TCC9	08tcc9 homo sapien
250	26	83.9	177	2	Q93CRO	Q93cro shigella bo	323	26	83.9	727	11	091VD9	091vd9 mus musculu
251	26	83.9	201	16	Q97JF7	Q97jf7 clostridium	324	26	83.9	729	5	Q9N4Y8	Q9n4y8 caenorhabdi
252	26	83.9	203	17	Q97ZT6	Q97zt6 sulfolobus	325	26	83.9	739	5	Q9V632	Q9v632 dirosophila
253	26	83.9	213	11	Q8VHH1	Q8vhh1 sulfolobus	326	26	83.9	751	5	Q9V6M9	Q9v6m9 dirosophila
254	26	83.9	214	2	Q8RR80	Q8rr80 notomys ale	327	26	83.9	757	5	Q9VYH2	Q9vyh2 dirosophila
255	26	83.9	221	17	Q9H1B3	Q9h1b3 acetobacter	328	26	83.9	760	10	023388	023388 arabidopsis
256	26	83.9	221	17	Q9H1B3	Q9h1b3 thermoplasm	329	26	83.9	762	2	Q9LCC0	Q9lcc0 alcaligenes
257	26	83.9	229	12	Q8TW5	Q8tw5 methanopyru	330	26	83.9	766	10	Q9ER95	Q9er95 thermoplasm
258	26	83.9	229	12	Q914G5	Q914g5 sulfolobus	331	26	83.9	780	17	Q9HJCS	Q9hjcs pseudomonas
259	26	83.9	242	5	Q9LSG8	Q9lsq8 arabidopsis	332	26	83.9	780	17	Q9HJCS	Q9hzx6 pseudomonas
260	26	83.9	249	5	Q9VBS3	Q9vbs3 arabidopsis	333	26	83.9	799	16	Q91J30	Q91j30 pseudomonas
261	26	83.9	253	16	Q970Z1	Q97bz3 dirosophila	334	26	83.9	830	16	Q9H2Y6	Q9h2y6 caenorhabdi
262	26	83.9	255	4	Q9P1A0	Q97gz1 streptococc	335	26	83.9	839	4	Q9UP83	Q9up83 homo sapien
263	26	83.9	259	12	Q8QZ08	Q931a0 homo sapien	336	26	83.9	849	10	Q9SHG6	Q9shg6 arabidopsis
264	26	83.9	261	16	P74291	Q8qz08 rana tigrin	337	26	83.9	886	13	Q42588	Q42588 xenopus lae
265	26	83.9	262	12	Q92547	P74291 synechocyst	338	26	83.9	904	11	008971	008971 mus musculu
266	26	83.9	275	10	Q9LDD8	Q92547 epizootic h	339	26	83.9	947	5	Q9B1A3	Q9bia3 caenorhabdi
267	26	83.9	281	16	Q8Y397	Q91d08 oryza sativ	340	26	83.9	975	10	Q9LMB4	Q9lmb4 arabidopsis
268	26	83.9	284	5	Q23758	Q8y397 raietonia s	341	26	83.9	975	16	Q8YK12	Q8yk12 anaabaena sp
269	26	83.9	284	5	Q17888	Q23758 clonorchis	342	26	83.9	990	2	Q9RPL8	Q9rpl8 streptococc
270	26	83.9	286	16	Q9AC04	Q17888 caenorhabdi	343	26	83.9	1054	10	Q9FKL3	Q9fkl3 arabidopsis
271	26	83.9	302	6	Q91416	Q9ac04 staphylococc	344	26	83.9	1057	11	Q91X04	Q91x04 mus musculu
272	26	83.9	303	6	Q95KC6	Q91416 pseudomonas	345	26	83.9	1139	10	Q8S665	Q8s665 oryza sativ
273	26	83.9	306	5	Q18585	Q95kc6 macaca fasc	346	26	83.9	1226	16	Q8Y745	Q8y745 bruceella me
274	26	83.9	308	17	Q8TN24	Q18585 caenorhabdi	347	26	83.9	1348	5	Q9VDT3	Q9vdt3 dirosophila
275	26	83.9	315	4	Q96JZ6	Q8tn24 methanosarc	348	26	83.9	1397	4	Q9NTG2	Q9ntg2 homo sapien
276	26	83.9	315	4	Q96S07	Q96jz6 homo sapien	349	26	83.9	1450	10	Q9SJR6	Q9sjr6 arabidopsis
277	26	83.9	315	6	Q9EE05	Q96s07 homo sapien	350	26	83.9	2468	5	Q9VEE5	Q9vee5 dirosophila
278	26	83.9	334	6	Q8RI75	Q9ee05 macaca fasc	351	26	83.9	2548	4	Q9UNJ2	Q9unj2 homo sapien
279	26	83.9	339	12	Q9YWT5	Q8ri75 fusobacteri	352	26	83.9	2626	11	Q9Z1N3	Q9z1n3 rattus norv
280	26	83.9	344	16	Q97SC7	Q9ywt5 melanoplus	353	26	83.9	3988	17	Q8TPE1	Q8tpel methanosarc
281	26	83.9	346	5	Q9Y6V9	Q97sc7 streptococc	354	26	83.9	4099	10	Q9C7Z6	Q9c7z6 arabidopsis
282	26	83.9	365	17	Q26986	Q9y6v9 dirosophila	355	26	83.9	4099	10	Q9C7Z6	Q9c7z6 arabidopsis
283	26	83.9	368	2	Q910S9	Q26986 methanobact	356	26	80.6	72	10	Q91XJ2	Q91xj2 arabidopsis
284	26	83.9	378	16	Q8UC02	Q91us9 rhizobium m	357	26	80.6	84	12	Q9DP94	Q9dp94 solanum tub
285	26	83.9	382	2	Q95380	Q8uc02 agrobacteri	358	26	80.6	85	12	Q82986	Q82986 kunjin viru
286	26	83.9	382	2	Q8RLK9	Q95380 escherichia	359	26	80.6	85	12	Q9DP94	Q9dp94 kunjin viru
287	26	83.9	384	16	Q9HTR9	Q8rlk9 salmonella	360	26	80.6	88	12	Q9DP94	Q9dp94 kunjin viru
288	26	83.9	402	4	Q9H766	Q9htr9 pseudomonas	361	26	80.6	92	15	Q76999	Q76999 human immun
289	26	83.9	408	3	Q94653	Q9h766 homo sapien	362	26	80.6	92	15	Q77000	Q77000 human immun
290	26	83.9	413	17	Q51984	Q94653 schizosacch	363	26	80.6	100	17	Q8TR13	Q8tr13 human immun
291	26	83.9	441	13	Q93403	Q51984 halobacteri	364	26	80.6	102	17	Q58485	Q58485 methanosarc
292	26	83.9	449	10	Q91S16	Q93403 lorpedo cal	365	26	80.6	104	2	Q9AMR2	Q9amr2 pyrococcus
293	26	83.9	449	13	Q93336	Q91s16 arabidopsis	366	26	80.6	114	16	Q8YDL3	Q8ydl3 bruceella me
294	26	83.9	458	17	Q59215	Q93336 gallus gall	367	26	80.6	117	9	Q38510	Q38510 bacterioplasm
295	26	83.9	470	5	Q9N8T1	Q59215 pyrococcus	368	26	80.6	117	9	Q38510	Q38510 bacterioplasm
296	26	83.9	476	13	Q93315	Q9n8t1 trypanosoma	369	26	80.6	117	9	Q38512	Q38512 bacterioplasm
297	26	83.9	485	17	Q9HLM6	Q93315 cyprinus ca	370	26	80.6	117	9	Q38512	Q38512 bacterioplasm
298	26	83.9	488	10	Q9LNG1	Q9hlm6 thermoplasm	371	26	80.6	117	9	Q38513	Q38513 bacterioplasm
299	26	83.9	498	11	Q64596	Q9lng1 arabidopsis	372	26	80.6	117	9	Q38513	Q38513 bacterioplasm
300	26	83.9	507	16	Q9POC4	Q64596 rattus norv	373	26	80.6	117	9	Q38515	Q38515 bacterioplasm
301	26	83.9	542	4	Q9H8T5	Q9poc4 ureaplasma	374	26	80.6	117	9	Q38515	Q38515 bacterioplasm
302	26	83.9	579	12	Q8OMW8	Q9h8t5 homo sapien	375	26	80.6	117	9	Q38517	Q38517 bacterioplasm
303	26	83.9	581	12	Q72760	Q8omw8 cowpox viru	376	26	80.6	117	9	Q38518	Q38518 bacterioplasm
304	26	83.9	596	16	Q97HM2	Q72760 cowpox viru	377	26	80.6	122	10	Q9SUU5	Q9suu5 arabidopsis
305	26	83.9	615	16	Q9PCG7	Q97hm2 clostridium	378	26	80.6	126	17	Q9HK12	Q9hk12 human immun
306	26	83.9	632	5	Q8SX70	Q9pcg7 xyliella fas	379	26	80.6	130	15	Q8UNP0	Q8unp0 human immun
307	26	83.9	637	17	Q972X6	Q8sx70 dirosophila	380	26	80.6	133	16	Q9HTR6	Q9htr6 pseudomonas
308	26	83.9	653	16	Q67097	Q972x6 sulfolobus	381	26	80.6	141	2	Q44784	Q44784 borrelia bu

382	25	80.6	141	2	Q9R3H3	Q9R3H3 borrella bu	455	25	80.6	372	16	Q67770	Q67770 aquifex aeo
383	25	80.6	141	2	Q9S0J3	Q9S0J3 borrella bu	456	25	80.6	374	16	Q8YRS1	Q8YRS1 anabaena sp
384	25	80.6	141	2	Q9S0A9	Q9S0A9 borrella bu	457	25	80.6	378	5	Q62012	Q62012 caenorhabdi
385	25	80.6	141	2	Q9S037	Q9S037 borrella bu	458	25	80.6	386	10	Q94776	Q94776 oryza sativ
386	25	80.6	141	2	Q9RXX5	Q9RXX5 borrella bu	459	25	80.6	388	13	Q90XK2	Q90XK2 brachydanio
387	25	80.6	141	2	Q07496	Q07496 borrella bu	460	25	80.6	390	5	Q9YPM4	Q9YPM4 drosophila
388	25	80.6	141	2	Q04478	Q04478 borrella bu	461	25	80.6	390	5	Q8SXR2	Q8SXR2 drosophila
389	25	80.6	141	2	Q04479	Q04479 borrella bu	462	25	80.6	392	3	Q9P883	Q9P883 aaracus di
390	25	80.6	141	2	Q9R9B8	Q9R9B8 borrella bu	463	25	80.6	393	2	Q9RQF1	Q9RQF1 mycoplasma
391	25	80.6	142	5	Q9VRC6	Q9VRC6 drosophila	464	25	80.6	394	16	Q97GV5	Q97GV5 clostridium
392	25	80.6	145	16	Q9H2T9	Q9H2T9 pseudomonas	465	25	80.6	397	5	Q8R0L0	Q8R0L0 drosophila
393	25	80.6	147	16	Q9X080	Q9X080 thermotoga	466	25	80.6	398	2	Q93P60	Q93P60 acroleplasm
394	25	80.6	162	16	Q94432	Q94432 ciona intes	467	25	80.6	401	11	Q91XC7	Q91XC7 mus muscu
395	25	80.6	162	16	Q04377	Q04377 bacillus su	468	25	80.6	402	16	Q8C106	Q8C106 lactococcus
396	25	80.6	163	16	Q92TW1	Q92TW1 rhizobium m	469	25	80.6	403	16	Q48345	Q48345 anabaena sp
397	25	80.6	183	16	Q9R5V3	Q9R5V3 thermoaer	470	25	80.6	402	16	Q97P00	Q97P00 streptococ
398	25	80.6	185	9	Q64022	Q64022 bacterioph	471	25	80.6	403	16	Q8YXO5	Q8YXO5 anabaena sp
399	25	80.6	185	16	Q03199	Q03199 bacillus su	472	25	80.6	403	16	Q8YNU6	Q8YNU6 anabaena sp
400	25	80.6	196	5	Q8SW34	Q8SW34 encephalit	473	25	80.6	403	16	Q8YKH3	Q8YKH3 anabaena sp
401	25	80.6	200	16	Q8UPX7	Q8UPX7 agrobacteri	474	25	80.6	410	11	Q9D5F1	Q9D5F1 mus muscu
402	25	80.6	203	16	Q8YBS4	Q8YBS4 bruceella me	475	25	80.6	413	16	Q8C177	Q8C177 pasteurella
403	25	80.6	208	4	Q75041	Q75041 homo sapien	476	25	80.6	413	16	Q8Y179	Q8Y179 anabaena sp
404	25	80.6	209	5	Q967H3	Q967H3 hydra atten	477	25	80.6	418	5	Q9VG86	Q9VG86 drosophila
405	25	80.6	209	12	Q80L21	Q80L21 sulfolobus	478	25	80.6	418	17	Q8TX44	Q8TX44 methanosarc
406	25	80.6	210	16	Q8Y1Y4	Q8Y1Y4 bruceella me	479	25	80.6	420	3	Q96U29	Q96U29 glibberella
407	25	80.6	216	4	Q96DH0	Q96DH0 homo sapien	480	25	80.6	420	11	Q9J310	Q9J310 mus muscu
408	25	80.6	222	12	Q9R0R6	Q9R0R6 myxoma viru	481	25	80.6	425	17	Q97W07	Q97W07 sulfolobus
409	25	80.6	234	5	Q9X1W7	Q9X1W7 entodinium	482	25	80.6	436	10	Q9SP99	Q9SP99 arabidopsis
410	25	80.6	235	16	Q928V7	Q928V7 listeria in	483	25	80.6	439	10	Q9XEP9	Q9XEP9 arabidopsis
411	25	80.6	238	17	Q928V7	Q928V7 archaeoglob	484	25	80.6	439	10	P92962	P92962 arabidopsis
412	25	80.6	241	16	Q97H70	Q97H70 clostridium	485	25	80.6	441	10	Q9XE48	Q9XE48 lycopersico
413	25	80.6	248	5	Q96945	Q96945 geodia cydo	486	25	80.6	442	10	Q9XE50	Q9XE50 lycopersico
414	25	80.6	254	11	Q9J120	Q9J120 mus muscu	487	25	80.6	442	10	P92961	P92961 arabidopsis
415	25	80.6	254	16	Q87599	Q87599 streptomyc	488	25	80.6	442	16	Q9PGK3	Q9PGK3 xylella fas
416	25	80.6	258	16	Q9X1Z8	Q9X1Z8 thermotoga	489	25	80.6	444	5	Q9YX43	Q9YX43 caenorhabdi
417	25	80.6	261	17	Q9V3X3	Q9V3X3 drosophila	490	25	80.6	450	10	Q94837	Q94837 hordium vul
418	25	80.6	261	17	Q9H1W9	Q9H1W9 thermoplas	491	25	80.6	450	16	Q91S64	Q91S64 pseudomonas
419	25	80.6	262	10	Q940K6	Q940K6 arabidopsis	492	25	80.6	453	10	Q91KX1	Q91KX1 atriplex ho
420	25	80.6	268	16	Q940T3	Q940T3 streptomyc	493	25	80.6	456	5	Q45301	Q45301 caenorhabdi
421	25	80.6	272	17	Q9H143	Q9H143 thermoplas	494	25	80.6	458	5	Q8SR56	Q8SR56 encephalit
422	25	80.6	274	16	Q9R3H2	Q9R3H2 rhizobium l	495	25	80.6	464	5	Q9U1R5	Q9U1R5 caenorhabdi
423	25	80.6	276	4	Q96F02	Q96F02 homo sapien	496	25	80.6	465	11	Q9D0L8	Q9D0L8 mus muscu
424	25	80.6	284	17	Q90202	Q90202 pyrococcus	497	25	80.6	467	5	Q9N8B5	Q9N8B5 trypanosoma
425	25	80.6	286	5	Q23011	Q23011 caenorhabdi	498	25	80.6	467	16	Q9K120	Q9K120 neisseria m
426	25	80.6	286	17	Q9YAJ9	Q9YAJ9 aeropyrum p	499	25	80.6	469	11	Q9N9B5	Q9N9B5 mus muscu
427	25	80.6	305	16	Q970Y0	Q970Y0 streptococ	500	25	80.6	469	5	Q8WPW7	Q8WPW7 paramicium
428	25	80.6	306	16	Q9HY52	Q9HY52 pseudomonas	501	25	80.6	471	16	Q8YBK0	Q8YBK0 bruceella me
429	25	80.6	313	2	Q9L8R2	Q9L8R2 pseudomonas	502	25	80.6	473	10	Q9LUN7	Q9LUN7 oryza sativ
430	25	80.6	316	5	Q9XMH8	Q9XMH8 caenorhabdi	503	25	80.6	476	4	Q9U1J9	Q9U1J9 homo sapien
431	25	80.6	316	5	Q9VHH0	Q9VHH0 drosophila	504	25	80.6	476	4	Q43148	Q43148 homo sapien
432	25	80.6	318	16	Q9U061	Q9U061 staphylococ	505	25	80.6	481	10	Q81498	Q81498 arabidopsis
433	25	80.6	326	15	Q9HTT2	Q9HTT2 pseudomonas	506	25	80.6	482	16	Q8XU27	Q8XU27 raietonia s
434	25	80.6	328	5	Q9V4H8	Q9V4H8 drosophila	507	25	80.6	487	10	Q92RC8	Q92RC8 carica papa
435	25	80.6	331	5	Q9XU69	Q9XU69 caenorhabdi	508	25	80.6	488	10	Q49819	Q49819 carica papa
436	25	80.6	331	10	Q9SXR1	Q9SXR1 arabidopsis	509	25	80.6	491	5	Q9GV11	Q9GV11 ephydalia f
437	25	80.6	332	2	Q69231	Q69231 bacillus sp	510	25	80.6	494	3	Q9C1E2	Q9C1E2 borrytis ci
438	25	80.6	333	10	Q9FFA1	Q9FFA1 arabidopsis	511	25	80.6	496	3	Q9HND0	Q9HND0 borrytis ci
439	25	80.6	333	16	Q9F3K0	Q9F3K0 streptomyc	512	25	80.6	498	3	Q9HND6	Q9HND6 borrytis ci
440	25	80.6	340	10	Q9FL46	Q9FL46 arabidopsis	513	25	80.6	498	3	Q9HNR1	Q9HNR1 borrytis ci
441	25	80.6	346	16	Q07414	Q07414 mycobacteri	514	25	80.6	498	3	Q9C1E3	Q9C1E3 borrytis ci
442	25	80.6	348	4	Q96G27	Q96G27 homo sapien	515	25	80.6	498	16	Q9RTW9	Q9RTW9 delinococcus
443	25	80.6	350	16	Q53543	Q53543 mycobacteri	516	25	80.6	500	3	Q9HNR5	Q9HNR5 borrytis ci
444	25	80.6	351	2	Q9RA92	Q9RA92 sinorhizobi	517	25	80.6	500	3	Q9C1E5	Q9C1E5 borrytis ci
445	25	80.6	351	17	Q97W55	Q97W55 sulfolobus	518	25	80.6	500	3	Q22904	Q22904 caenorhabdi
446	25	80.6	357	16	Q8UDH4	Q8UDH4 agrobacteri	519	25	80.6	501	3	Q9HNR4	Q9HNR4 borrytis ci
447	25	80.6	359	16	Q8RC28	Q8RC28 fusobacteri	520	25	80.6	501	3	Q9HNR3	Q9HNR3 borrytis ci
448	25	80.6	363	16	P73886	P73886 synochocyst	521	25	80.6	501	3	Q9HDO9	Q9HDO9 borrytis ci
449	25	80.6	365	16	P73756	P73756 synochocyst	522	25	80.6	501	3	Q9C1E4	Q9C1E4 borrytis ci
450	25	80.6	367	10	Q9M6B1	Q9M6B1 chlamydomon	523	25	80.6	502	3	Q9C1E1	Q9C1E1 borrytis ci
451	25	80.6	367	10	Q9M6B1	Q9M6B0 chlamydomon	524	25	80.6	503	3	Q9HNR2	Q9HNR2 borrytis ci
452	25	80.6	369	3	Q96W00	Q96W00 geotrichum	525	25	80.6	504	3	Q9C1E6	Q9C1E6 borrytis ci
453	25	80.6	370	17	Q29775	Q29775 archaeoglob	526	25	80.6	504	4	Q94996	Q94996 homo sapien
454	25	80.6	371	12	Q69294	Q69294 gallid hep	527	25	80.6	505	5	Q9YVH4	Q9YVH4 drosophila

528	25	80.6	509	5	Q95W10	Q95W10 drosophila	601	25	80.6	970	12	Q80Z06	Q80Z06 rana tigris
529	25	80.6	512	15	Q9W14	Q9W14 drosophila	602	25	80.6	971	16	Q9ZE42	Q9ZE42 rickettsia
530	25	80.6	512	15	Q8XYK0	Q8XYK0 ralstonia s	603	25	80.6	976	16	Q97EV0	Q97EV0 clostridium
531	25	80.6	517	12	Q9YU5L	Q9YU5L turkey aden	604	25	80.6	979	16	Q98PK8	Q98PK8 mycoplasma
532	25	80.6	522	3	Q93867	Q93867 erysiphe gr	605	25	80.6	983	10	Q9EP33	Q9EP33 oryza sativ
533	25	80.6	522	3	Q9P428	Q9P428 bettyrcis ci	606	25	80.6	1000	17	Q96YK3	Q96YK3 sulfolobus
534	25	80.6	522	3	Q8TGV1	Q8TGV1 monilia f	607	25	80.6	1058	5	Q96J35	Q96J35 drosophila
535	25	80.6	525	5	Q9VW60	Q9VW60 drosophila	608	25	80.6	1075	5	Q95U75	Q95U75 drosophila
536	25	80.6	534	5	Q21063	Q21063 caenorhabdi	609	25	80.6	1115	2	Q93R34	Q93R34 streptococ
537	25	80.6	535	10	Q9S127	Q9S127 arabadopsi	610	25	80.6	1118	4	Q9UIM0	Q9UIM0 homo sapien
538	25	80.6	536	17	Q97YB2	Q97YB2 sulfolobus	611	25	80.6	1133	10	Q82504	Q82504 arabadopsi
539	25	80.6	536	17	Q96Z08	Q96Z08 sulfolobus	612	25	80.6	1185	10	Q9AY94	Q9AY94 oryza sativ
540	25	80.6	539	5	Q9YK44	Q9YK44 drosophila	613	25	80.6	1205	5	Q96694	Q96694 plasmodium
541	25	80.6	542	2	Q96294	Q96294 thibacilli	614	25	80.6	1255	5	Q8SVZ5	Q8SVZ5 encephalito
542	25	80.6	548	2	Q86288	Q86288 lactococcus	615	25	80.6	1257	12	Q8QYV6	Q8QYV6 euprosterna
543	25	80.6	549	10	Q93239	Q93239 arabadopsi	616	25	80.6	1328	6	Q97836	Q97836 syncerus ca
544	25	80.6	550	16	Q25060	Q25060 helicobacte	617	25	80.6	1332	6	Q97897	Q97897 tragelaphus
545	25	80.6	550	16	Q9ZME9	Q9ZME9 helicobacte	618	25	80.6	1332	6	Q95325	Q95325 bos taurus
546	25	80.6	556	16	Q8XZ51	Q8XZ51 ralstonia s	619	25	80.6	1401	10	Q94LJ6	Q94LJ6 oryza sativ
547	25	80.6	558	5	Q19290	Q19290 caenorhabdi	620	25	80.6	1416	10	Q8S609	Q8S609 oryza sativ
548	25	80.6	568	5	Q9V903	Q9V903 drosophila	621	25	80.6	1450	10	Q9LPH0	Q9LPH0 xenopus lae
549	25	80.6	575	16	Q8YK47	Q8YK47 anabaena sp	622	25	80.6	1490	13	P79522	P79522 aspergillus
550	25	80.6	577	5	Q9HFR9	Q9HFR9 escherichia	623	25	80.6	1497	10	Q8S718	Q8S718 oryza sativ
551	25	80.6	584	5	Q25283	Q25283 leishmania	624	25	80.6	1529	3	Q9Y7C6	Q9Y7C6 aspergillus
552	25	80.6	588	16	Q99UV8	Q99UV8 staphylococ	625	25	80.6	1676	10	Q23332	Q23332 arabadopsi
553	25	80.6	589	5	Q19289	Q19289 caenorhabdi	626	25	80.6	1680	10	Q64588	Q64588 arabadopsi
554	25	80.6	594	4	Q96HN9	Q96HN9 homo sapien	627	25	80.6	1720	10	Q9AV68	Q9AV68 oryza sativ
555	25	80.6	595	10	Q9XEX0	Q9XEX0 nicotiana t	628	25	80.6	1769	5	Q16625	Q16625 caenorhabdi
556	25	80.6	595	16	Q9A9X2	Q9A9X2 caulobacter	629	25	80.6	2183	12	Q98333	Q98333 measles vir
557	25	80.6	603	10	Q9S2T9	Q9S2T9 arabadopsi	630	25	80.6	2233	12	Q9J7D6	Q9J7D6 bovine para
558	25	80.6	603	11	Q9UJ03	Q9UJ03 mus musculu	631	25	80.6	2233	12	Q9J7D6	Q9J7D6 bovine para
559	25	80.6	607	12	Q8Q5A2	Q8Q5A2 felis domes	632	25	80.6	2233	12	Q81890	Q81890 human para
560	25	80.6	609	5	Q962H9	Q962H9 toxoplasma	633	25	80.6	2258	12	Q81081	Q81081 human para
561	25	80.6	614	10	Q94DX2	Q94DX2 oryza sativ	634	25	80.6	2556	12	Q9CH56	Q9CH56 gallid hep
562	25	80.6	615	5	Q961X3	Q961X3 drosophila	635	25	80.6	2556	12	Q9CH56	Q9CH56 gallid hep
563	25	80.6	617	17	Q9HKM1	Q9HKM1 thermoplas	636	25	80.6	7576	2	Q9KTE0	Q9KTE0 streptomyce
564	25	80.6	618	5	Q8TMS2	Q8TMS2 encephalito	637	25	80.6	7576	2	Q9ZG44	Q9ZG44 streptomyce
565	25	80.6	632	10	Q8W3I6	Q8W3I6 oryza sativ	638	25	80.6	7576	2	Q9ZG44	Q9ZG44 streptomyce
566	25	80.6	633	5	Q9VZ84	Q9VZ84 drosophila	639	25	80.6	77.4	42	Q9H453	Q9H453 homo sapien
567	25	80.6	643	16	Q8XKX0	Q8XKX0 clostridium	640	25	80.6	77.4	42	Q9YCT4	Q9YCT4 aerypyrum p
568	25	80.6	654	10	Q8S7H9	Q8S7H9 oryza sativ	641	25	80.6	77.4	55	Q9GNZ8	Q9GNZ8 plasmodium
569	25	80.6	664	10	Q9STY9	Q9STY9 arabadopsi	642	25	80.6	77.4	55	Q9GZ89	Q9GZ89 plasmodium
570	25	80.6	669	17	Q97CU9	Q97CU9 thermoplas	643	25	80.6	77.4	55	Q9GZ83	Q9GZ83 plasmodium
571	25	80.6	671	5	Q9WYV6	Q9WYV6 penaeus van	644	25	80.6	77.4	55	Q9HNV7	Q9HNV7 japonese en
572	25	80.6	678	12	Q8V3A7	Q8V3A7 caspalia e	645	25	80.6	77.4	58	Q9GNE0	Q9GNE0 plasmodium
573	25	80.6	678	12	Q8TMR7	Q8TMR7 methanosarc	646	25	80.6	77.4	58	Q9GNB9	Q9GNB9 plasmodium
574	25	80.6	689	16	Q9KY46	Q9KY46 streptomyce	647	25	80.6	77.4	58	Q9GN91	Q9GN91 plasmodium
575	25	80.6	696	5	Q8SR54	Q8SR54 encephalito	648	25	80.6	77.4	58	Q9GN77	Q9GN77 plasmodium
576	25	80.6	698	11	Q8R3J2	Q8R3J2 mus musculu	649	25	80.6	77.4	58	Q9GN76	Q9GN76 plasmodium
577	25	80.6	699	5	Q95ZC2	Q95ZC2 leishmania	650	25	80.6	77.4	58	Q9GZB7	Q9GZB7 plasmodium
578	25	80.6	699	11	Q9DA66	Q9DA66 mus musculu	651	25	80.6	77.4	58	Q9GZB7	Q9GZB7 plasmodium
579	25	80.6	730	12	Q91TML	Q91TML tupala hep	652	25	80.6	77.4	58	Q9GZB6	Q9GZB6 plasmodium
580	25	80.6	735	10	Q93Z03	Q93Z03 arabadopsi	653	25	80.6	77.4	58	Q9GZB2	Q9GZB2 plasmodium
581	25	80.6	736	10	Q82754	Q82754 arabadopsi	654	25	80.6	77.4	58	Q9GZB1	Q9GZB1 plasmodium
582	25	80.6	754	4	Q9P0L1	Q9P0L1 homo sapien	655	25	80.6	77.4	58	Q9GZA9	Q9GZA9 plasmodium
583	25	80.6	754	17	Q8TPU1	Q8TPU1 methanosarc	656	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
584	25	80.6	778	5	Q9BLD4	Q9BLD4 palinopecte	657	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
585	25	80.6	785	10	Q9FRM0	Q9FRM0 oryza sativ	658	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
586	25	80.6	786	16	Q9FBS2	Q9FBS2 streptomyce	659	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
587	25	80.6	787	17	Q9UYH2	Q9UYH2 pyrococcus	660	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
588	25	80.6	801	16	Q8YX11	Q8YX11 anabaena sp	661	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
589	25	80.6	806	5	Q8J2A3	Q8J2A3 dictyosteli	662	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
590	25	80.6	832	5	Q963E6	Q963E6 drosophila	663	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
591	25	80.6	847	16	Q92KG2	Q92KG2 rhizobium m	664	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
592	25	80.6	847	16	Q99P14	Q99P14 mus musculu	665	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
593	25	80.6	866	16	Q8UC06	Q8UC06 agrobacteri	666	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
594	25	80.6	869	2	Q9R912	Q9R912 rhizobium m	667	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
595	25	80.6	884	9	Q9T145	Q9T145 bacterioph	668	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
596	25	80.6	897	5	Q9XZRO	Q9XZRO caenorhabdi	669	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
597	25	80.6	901	5	Q9TX17	Q9TX17 caenorhabdi	670	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
598	25	80.6	929	16	Q86810	Q86810 streptomyce	671	25	80.6	77.4	58	Q9GZAS	Q9GZAS plasmodium
599	25	80.6	938	5	Q9V8V8	Q9V8V8 drosophila	672	25	80.6	77.4	64	Q72732	Q72732 cowpox viru
600	25	80.6	954	5	Q9U0MS	Q9U0MS plasmodium	673	25	80.6	77.4	65	Q82900	Q82900 japonese en

674	24	77.4	66	5	Q9HAP6	Q9HAP6 drosophila	747	24	77.4	164	16	Q9PNA4	Q9PNA4 campylobact
675	24	77.4	68	12	Q82899	Q82899 japanese en	748	24	77.4	165	15	Q91204	Q91204 pseudomonas
676	24	77.4	69	17	Q8TQC7	Q8TQC7 methanosarc	749	24	77.4	165	17	Q98810	Q98810 atcheaglob
677	24	77.4	82	10	Q9L146	Q9L146 phaseolus v	750	24	77.4	166	10	Q04125	Q04125 nicotiana t
678	24	77.4	84	12	Q82898	Q82898 japanese en	751	24	77.4	167	2	Q8V1N8	Q8V1N8 pseudomonas
679	24	77.4	86	10	Q24439	Q24439 phaseolus v	752	24	77.4	169	5	Q20131	Q20131 caenorhabdi
680	24	77.4	88	12	Q9WPH6	Q9WPH6 japanese en	753	24	77.4	170	2	Q46103	Q46103 campylobact
681	24	77.4	89	16	Q9KUT7	Q9KUT7 vibrio chol	754	24	77.4	171	10	Q9FGY5	Q9FGY5 arabisdopsis
682	24	77.4	90	12	Q9MNV8	Q9MNV8 japanese en	755	24	77.4	171	10	Q92SW5	Q92SW5 hamamelis v
683	24	77.4	91	16	Q9KVV0	Q9KVV0 vibrio chol	756	24	77.4	172	10	Q4668	Q4668 theobroma c
684	24	77.4	94	2	Q32635	Q32635 helicobacte	757	24	77.4	173	10	Q946A1	Q946A1 theobroma c
685	24	77.4	95	10	Q9S7N8	Q9S7N8 glycine max	758	24	77.4	174	10	Q945X7	Q945X7 theobroma c
686	24	77.4	100	16	Q931X7	Q931X7 staphylococ	759	24	77.4	174	17	Q9V161	Q9V161 pyrococcus
687	24	77.4	102	13	Q91662	Q91662 xenopus lae	760	24	77.4	178	17	Q975A2	Q975A2 sulfolobus
688	24	77.4	110	16	Q8XSH1	Q8XSH1 escherichia	761	24	77.4	180	10	Q9STB9	Q9STB9 gracillariop
689	24	77.4	111	12	Q9DKM7	Q9DKM7 spodiopetra	762	24	77.4	183	10	Q9SHU8	Q9SHU8 arabisdopsis
690	24	77.4	111	17	Q976V1	Q976V1 sulfolobus	763	24	77.4	184	12	Q2YV50	Q2YV50 melanoplus
691	24	77.4	112	2	Q9L484	Q9L484 uncultured	764	24	77.4	185	16	Q67590	Q67590 aquifex aeo
692	24	77.4	112	2	Q9L483	Q9L483 uncultured	765	24	77.4	187	4	Q96Q77	Q96Q77 homo sapien
693	24	77.4	112	2	Q9L480	Q9L480 uncultured	766	24	77.4	189	8	Q9NCAS	Q9NCAS nephroselm
694	24	77.4	112	2	Q9L478	Q9L478 uncultured	767	24	77.4	190	8	Q9MD76	Q9MD76 lupinus lut
695	24	77.4	112	2	Q9L477	Q9L477 uncultured	768	24	77.4	190	8	Q9MD58	Q9MD58 lupinus ang
696	24	77.4	114	5	Q22333	Q22333 caenorhabdi	769	24	77.4	190	8	Q9MFG5	Q9MFG5 lupinus ang
697	24	77.4	114	16	Q8X3N5	Q8X3N5 escherichia	770	24	77.4	190	8	Q9MFG7	Q9MFG7 lupinus mut
698	24	77.4	115	16	Q53334	Q53334 mycobacteri	771	24	77.4	190	8	Q9MFG6	Q9MFG6 lupinus alb
699	24	77.4	117	2	Q9L4C2	Q9L4C2 uncultured	772	24	77.4	190	16	Q9MW48	Q9MW48 pseudomonas
700	24	77.4	117	2	Q9L4C1	Q9L4C1 uncultured	773	24	77.4	192	8	Q9MFE5	Q9MFE5 beta vulgar
701	24	77.4	117	2	Q9L4C0	Q9L4C0 uncultured	774	24	77.4	192	12	Q9Q3W5	Q9Q3W5 hepatitis c
702	24	77.4	117	2	Q9L4B9	Q9L4B9 uncultured	775	24	77.4	193	16	Q8XE96	Q8XE96 escherichia
703	24	77.4	117	2	Q9L4B8	Q9L4B8 uncultured	776	24	77.4	193	17	Q97U03	Q97U03 sulfolobus
704	24	77.4	117	2	Q9L4B6	Q9L4B6 uncultured	777	24	77.4	194	10	Q9LZ13	Q9LZ13 arabisdopsis
705	24	77.4	117	2	Q9L4B5	Q9L4B5 uncultured	778	24	77.4	197	16	Q9A6X2	Q9A6X2 caulobacter
706	24	77.4	117	2	Q9L4B4	Q9L4B4 uncultured	779	24	77.4	199	16	P73231	P73231 synechocyst
707	24	77.4	117	2	Q9L4B3	Q9L4B3 uncultured	780	24	77.4	199	17	Q6U385	Q6U385 pyrococcus
708	24	77.4	117	2	Q9L4B1	Q9L4B1 uncultured	781	24	77.4	200	8	Q98RNO	Q98RNO guillardia
709	24	77.4	117	2	Q9L4A9	Q9L4A9 uncultured	782	24	77.4	200	15	Q99ZU2	Q99ZU2 streptococ
710	24	77.4	117	2	Q9L4A8	Q9L4A8 uncultured	783	24	77.4	200	15	Q8U862	Q8U862 agrobacteri
711	24	77.4	117	2	Q9L4A7	Q9L4A7 uncultured	784	24	77.4	201	16	Q67778	Q67778 aquifex aeo
712	24	77.4	117	2	Q9L4A6	Q9L4A6 uncultured	785	24	77.4	201	16	Q96KQ7	Q96KQ7 rhlzodum l
713	24	77.4	117	2	Q9L4A2	Q9L4A2 uncultured	786	24	77.4	203	16	Q34487	Q34487 bacillus su
714	24	77.4	117	2	Q9L4A1	Q9L4A1 uncultured	787	24	77.4	204	5	Q26014	Q26014 plasmodium
715	24	77.4	117	2	Q9L4A0	Q9L4A0 uncultured	788	24	77.4	206	8	Q954U3	Q954U3 beta vulgar
716	24	77.4	117	2	Q9L4E9	Q9L4E9 uncultured	789	24	77.4	208	17	Q9UXD8	Q9UXD8 sulfolobus
717	24	77.4	117	2	Q9L4A9	Q9L4A9 uncultured	790	24	77.4	210	16	Q97PK9	Q97PK9 clostridium
718	24	77.4	117	5	Q9VYX6	Q9VYX6 drosophila	791	24	77.4	215	2	Q9XDR3	Q9XDR3 clostridium
719	24	77.4	117	17	Q97UX4	Q97UX4 sulfolobus	792	24	77.4	215	16	Q8X6R8	Q8X6R8 escherichia
720	24	77.4	121	2	Q48516	Q48516 leptospira	793	24	77.4	216	16	Q85823	Q85823 yersinia pe
721	24	77.4	125	4	Q12835	Q12835 homo sapien	794	24	77.4	216	17	Q8ZWP8	Q8ZWP8 pyrobaculum
722	24	77.4	126	16	Q8UD18	Q8UD18 agrobacteri	795	24	77.4	218	16	Q8ZD58	Q8ZD58 yersinia pe
723	24	77.4	127	16	Q8YR79	Q8YR79 anabaena sp	796	24	77.4	219	17	Q8U470	Q8U470 pyrococcus
724	24	77.4	127	16	Q9ZNG7	Q9ZNG7 staphylococ	797	24	77.4	220	16	Q9C1M2	Q9C1M2 pasteurella
725	24	77.4	130	4	Q14942	Q14942 homo sapien	798	24	77.4	222	17	Q975F4	Q975F4 sulfolobus
726	24	77.4	130	17	Q9YFC0	Q9YFC0 aeropyrum p	799	24	77.4	223	17	Q8TVX5	Q8TVX5 methanopyru
727	24	77.4	131	16	Q9ZCE2	Q9ZCE2 listeria in	800	24	77.4	225	16	Q8X1M6	Q8X1M6 clostridium
728	24	77.4	132	17	Q9HBP1	Q9HBP1 halobacteri	801	24	77.4	226	13	Q9PW60	Q9PW60 cyprinus ca
729	24	77.4	134	16	Q8R7K1	Q8R7K1 thermoaer	802	24	77.4	227	16	Q9FE23	Q9FE23 streptomyce
730	24	77.4	137	10	Q9LPP4	Q9LPP4 arabisdopsis	803	24	77.4	229	10	Q43363	Q43363 nicotiana t
731	24	77.4	137	12	Q9PXM1	Q9PXM1 hepatitis c	804	24	77.4	230	2	Q931F2	Q931F2 propionibac
732	24	77.4	139	2	Q50234	Q50234 zymomonas m	805	24	77.4	230	13	Q8QGE7	Q8QGE7 coturnix co
733	24	77.4	141	12	Q9YW64	Q9YW64 melanoplus	806	24	77.4	231	5	Q9U8G2	Q9U8G2 trypanosoma
734	24	77.4	142	16	Q9C1Y8	Q9C1Y8 iactococcus	807	24	77.4	231	16	Q53480	Q53480 mycobacteri
735	24	77.4	145	12	Q99ES0	Q99ES0 japanese en	808	24	77.4	232	16	Q8Z1D1	Q8Z1D1 yersinia pe
736	24	77.4	147	13	Q12940	Q12940 gallus gall	809	24	77.4	233	16	Q8Z4V0	Q8Z4V0 listeria mo
737	24	77.4	150	3	P74980	P74980 yersinia ps	810	24	77.4	233	16	Q8YCB2	Q8YCB2 brucella me
738	24	77.4	150	3	Q14087	Q14087 schistosach	811	24	77.4	235	5	Q9NEM6	Q9NEM6 caenorhabdi
739	24	77.4	150	16	Q9ZEPI	Q9ZEPI streptomyce	812	24	77.4	236	4	Q9NMG0	Q9NMG0 homo sapien
740	24	77.4	151	16	Q99RH4	Q99RH4 staphylococ	813	24	77.4	237	16	Q8YVW8	Q8YVW8 anabaena sp
741	24	77.4	152	10	Q01920	Q01920 phytophthor	814	24	77.4	238	5	Q9B1Y9	Q9B1Y9 renilla mu
742	24	77.4	152	10	Q01918	Q01918 phytophthor	815	24	77.4	238	16	Q67863	Q67863 aquifex aeo
743	24	77.4	156	16	Q8RP01	Q8RP01 thermoaer	816	24	77.4	241	16	Q9KB70	Q9KB70 bacillus ha
744	24	77.4	157	5	Q9U7E2	Q9U7E2 pecten maxi	817	24	77.4	242	16	Q9ZMU1	Q9ZMU1 helicobacte
745	24	77.4	158	10	Q9AS17	Q9AS17 caulobacter	818	24	77.4	243	16	Q24951	Q24951 helicobacte
746	24	77.4	160	16	Q9MAW2	Q9MAW2 brugutera g	819	24	77.4	243	3	Q12134	Q12134 saccharomyc

820	24	77.4	243	9	064333	064333 bacterioph	893	24	77.4	297	16	050726	050726 borrelia bu
821	24	77.4	244	16	08X3V9	08X3V9 listeria mo	894	24	77.4	297	16	050805	050805 borrelia bu
822	24	77.4	248	10	022210	022210 arabidopsis	895	24	77.4	300	13	09P84	09P84 carassius a
823	24	77.4	248	16	09KCS1	09KCS1 bacillus ha	896	24	77.4	300	16	09WX26	09WX26 streptomye
824	24	77.4	248	16	097DB9	097DB9 clostridium	897	24	77.4	301	10	09F539	09F539 zea mays (m
825	24	77.4	249	16	09K710	09K710 bacillus ha	898	24	77.4	301	16	09KX28	09KX28 vibrio chol
826	24	77.4	251	5	09NKM6	09NKM6 leishmania	899	24	77.4	302	10	041777	041777 zea mays (m
827	24	77.4	251	17	029354	029354 archaeoglob	900	24	77.4	303	4	09H317	09H317 homo sapien
828	24	77.4	252	10	040701	040701 oryza sativ	901	24	77.4	303	5	017706	017706 caenorhabd
829	24	77.4	253	16	08XV4	08XV4 escherichia	902	24	77.4	304	5	08SUA7	08SUA7 encephalito
830	24	77.4	254	10	065103	065103 oryza sativ	903	24	77.4	304	10	041776	041776 zea mays (m
831	24	77.4	254	11	09CT99	09CT99 mus musculu	904	24	77.4	304	10	08S9W8	08S9W8 oryza sativ
832	24	77.4	254	17	097M63	097M63 sulfolobus	905	24	77.4	309	2	09FDP8	09FDP8 staphylococ
833	24	77.4	255	16	09C118	09C118 lactococcus	906	24	77.4	309	16	024916	024916 helicobacte
834	24	77.4	255	16	08ZB31	08ZB31 yersinia pe	907	24	77.4	309	16	09ZMY1	09ZMY1 helicobacte
835	24	77.4	255	16	08RBP5	08RBP5 thermoaer	908	24	77.4	310	17	08U307	08U307 pyrococcus
836	24	77.4	258	2	06E280	06E280 escherichia	909	24	77.4	312	2	09L796	09L796 portnyromon
837	24	77.4	258	5	09GQ17	09GQ17 caenorhabd	910	24	77.4	312	4	09S286	09S286 homo sapien
838	24	77.4	259	16	09RXU2	09RXU2 deinococcus	911	24	77.4	312	16	098QY0	098QY0 mycoplasma
839	24	77.4	260	5	017343	017343 caenorhabd	912	24	77.4	315	2	09XWJ3	09XWJ3 pseudomonas
840	24	77.4	264	4	096AS3	096AS3 homo sapien	913	24	77.4	319	16	09S2H1	09S2H1 streptomye
841	24	77.4	265	2	09KW39	09KW39 wolbachia's	914	24	77.4	319	2	093PS4	093PS4 comamonas t
842	24	77.4	265	10	004126	004126 nicotiana t	915	24	77.4	319	16	067018	067018 aquifex aeo
843	24	77.4	268	10	040457	040457 streptococ	916	24	77.4	319	17	097U08	097U08 sulfobus
844	24	77.4	268	2	082875	082875 streptococ	917	24	77.4	320	16	09A7Z8	09A7Z8 caulobacter
845	24	77.4	269	3	002392	002392 saccharomyc	918	24	77.4	321	11	09CXE2	09CXE2 mus musculu
846	24	77.4	269	5	018670	018670 drosophila	919	24	77.4	321	11	09CXE2	09CXE2 mus musculu
847	24	77.4	269	5	016076	016076 drosophila	920	24	77.4	322	12	09JPE8	09JPE8 vaccinia vl
848	24	77.4	269	5	016087	016087 drosophila	921	24	77.4	323	12	087633	087633 cowpox viru
849	24	77.4	269	10	09SKW3	09SKW3 arabidopsis	922	24	77.4	323	12	08QX72	08QX72 camelpox vl
850	24	77.4	270	16	08RE14	08RE14 fusobacteri	923	24	77.4	323	12	08QX13	08QX13 cowpox viru
851	24	77.4	271	2	047308	047308 escherichia	924	24	77.4	326	2	093653	093653 serattia ma
852	24	77.4	271	2	09K3D9	09K3D9 uncultured	925	24	77.4	327	16	09G870	09G870 caenorhabd
853	24	77.4	271	2	054477	054477 escherichia	926	24	77.4	327	5	09U2B5	09U2B5 streptomye
854	24	77.4	271	10	005786	005786 klebsiella	927	24	77.4	328	16	08X115	08X115 salmonella
855	24	77.4	273	10	09STR2	09STR2 arabidopsis	928	24	77.4	329	2	09F6F5	09F6F5 bacteroides
856	24	77.4	273	16	08XU19	08XU19 clostridium	929	24	77.4	330	5	09NAP9	09NAP9 caenorhabd
857	24	77.4	274	16	08RNV2	08RNV2 pseudomonas	930	24	77.4	331	10	09FX56	09FX56 arabidopsis
858	24	77.4	275	10	09MSA7	09MSA7 zea mays (m	931	24	77.4	331	10	09AM15	09AM15 guillardia
859	24	77.4	275	16	08RBU0	08RBU0 thermoaer	932	24	77.4	332	2	030670	030670 bacillus su
860	24	77.4	277	16	09CF12	09CF12 lactococcus	933	24	77.4	333	5	09BJ52	09BJ52 heterodera
861	24	77.4	278	10	041851	041851 zea mays (m	934	24	77.4	334	16	08YPR9	08YPR9 anaplasma sp
862	24	77.4	279	16	055362	055362 synechocyst	935	24	77.4	335	5	09BLJ5	09BLJ5 acrythosiph
863	24	77.4	279	16	08XWS9	08XWS9 clostridium	936	24	77.4	335	16	09RUP4	09RUP4 chlamydia m
864	24	77.4	281	5	09UTAB	09UTAB brachionus	937	24	77.4	336	5	09USE5	09USE5 entamoeba d
865	24	77.4	281	10	08RWD7	08RWD7 arabidopsis	938	24	77.4	336	16	09BP68	09BP68 thizobium l
866	24	77.4	281	17	0979W3	0979W3 thermoplasm	939	24	77.4	336	16	09WZ28	09WZ28 thermocoga
867	24	77.4	282	16	09KU52	09KU52 vibrio chol	940	24	77.4	339	17	058916	058916 pyrococcus
868	24	77.4	283	6	09BGG0	09BGG0 mustela put	941	24	77.4	339	17	08RHI3	08RHI3 methanosarc
869	24	77.4	284	17	09HUR5	09HUR5 thermoplasm	942	24	77.4	340	12	09YYO6	09YYO6 melanoplus
870	24	77.4	285	5	026013	026013 plasmodium	943	24	77.4	342	5	09SV86	09SV86 nosema locu
871	24	77.4	287	8	09S980	09S980 triticum ae	944	24	77.4	343	12	08V543	08V543 monkeypox v
872	24	77.4	290	4	09BSRO	09BSRO homo sapien	945	24	77.4	343	16	08ZP83	08ZP83 salmonella
873	24	77.4	290	5	09UTB4	09UTB4 adineta vag	946	24	77.4	343	16	08Z789	08Z789 salmonella
874	24	77.4	290	5	09NBE3	09NBE3 adineta vag	947	24	77.4	345	17	09HNM8	09HNM8 halobacteri
875	24	77.4	290	5	09NBE2	09NBE2 adineta vag	948	24	77.4	346	5	016328	016328 caenorhabd
876	24	77.4	291	5	09U7B3	09U7B3 habroclitrocha	949	24	77.4	347	16	099WX3	099WX3 staphylococ
877	24	77.4	291	5	09U7B2	09U7B2 philodina r	950	24	77.4	348	12	089110	089110 variola vir
878	24	77.4	291	5	09U7A5	09U7A5 esophora eh	951	24	77.4	348	12	089754	089754 variola vir
879	24	77.4	291	5	09NBE6	09NBE6 philodina r	952	24	77.4	348	12	089088	089088 variola vir
880	24	77.4	291	5	09NBE5	09NBE5 habroclitrocha	953	24	77.4	349	12	09ZEF4	09ZEF4 listeria in
881	24	77.4	291	5	09NBE4	09NBE4 habroclitrocha	954	24	77.4	349	16	08VX27	08VX27 camelpox vl
882	24	77.4	291	16	P72937	P72937 synechocyst	955	24	77.4	350	10	09CC9C	09CC9C arabidopsis
883	24	77.4	291	12	097Z88	097Z88 sulfolobus	956	24	77.4	350	10	094TW2	094TW2 oryza sativ
884	24	77.4	292	12	036377	036377 alcelaphine	957	24	77.4	351	2	09ADY7	09ADY7 agrobacteri
885	24	77.4	292	12	055748	055748 chilo iride	958	24	77.4	351	10	09FKD1	09FKD1 arabidopsis
886	24	77.4	293	5	09MBE1	09MBE1 philodina r	959	24	77.4	351	10	09EX55	09EX55 arabidopsis
887	24	77.4	293	5	09NBE9	09NBE9 philodina r	960	24	77.4	351	10	09S733	09S733 arabidopsis
888	24	77.4	295	16	09RTR9	09RTR9 deinococcus	961	24	77.4	352	4	09GJ72	09GJ72 homo sapien
889	24	77.4	296	5	09U7A7	09U7A7 brachionus	962	24	77.4	354	2	059329	059329 comamonas t
890	24	77.4	296	5	09U7A6	09U7A6 seison neba	963	24	77.4	354	12	089406	089406 paramonacium
891	24	77.4	296	17	027257	027257 methanobact	964	24	77.4	355	2	005366	005366 actinobacil
892	24	77.4	297	4	09FOR4	09FOR4 homo sapien	965	24	77.4	355	2	066249	066249 actinobacil

966	24	77.4	355	2	Q9XD08	Q9XD08 actinobacill
967	24	77.4	355	2	Q9JRS	Q9JRS actinobacill
968	24	77.4	355	2	Q9AQC0	Q9AQC0 actinobacill
969	24	77.4	355	16	Q8UC06	Q8UC06 agrobacteri
970	24	77.4	356	16	Q99WH9	Q99WH9 staphylococ
971	24	77.4	356	16	Q99WH5	Q99WH5 staphylococ
972	24	77.4	356	5	Q9NRS	Q9NRS drosophila
973	24	77.4	359	5	Q9NRS	Q9NRS drosophila
974	24	77.4	359	5	Q9NRS	Q9NRS drosophila
975	24	77.4	360	5	Q20723	Q20723 caenorhabd
976	24	77.4	361	2	Q85776	Q85776 rhizobium l
977	24	77.4	361	16	Q9KBP5	Q9KBP5 bacillus ha
978	24	77.4	362	2	Q9S162	Q9S162 erysipelo
979	24	77.4	362	2	Q9S160	Q9S160 erysipelo
980	24	77.4	365	8	Q32831	Q32831 pyramonas
981	24	77.4	365	2	Q9R6W8	Q9R6W8 anabaena sp
982	24	77.4	365	5	Q9XTE4	Q9XTE4 caenorhabd
983	24	77.4	366	2	Q9X406	Q9X406 methylosulf
984	24	77.4	366	16	Q51031	Q51031 borrelia bu
985	24	77.4	367	5	Q17575	Q17575 caenorhabd
986	24	77.4	368	17	Q8ZVP6	Q8ZVP6 pyrobaculum
987	24	77.4	370	5	Q8SR44	Q8SR44 encephalit
988	24	77.4	370	16	Q66865	Q66865 aquilex aeo
989	24	77.4	372	4	Q15391	Q15391 homo sapien
990	24	77.4	372	4	Q96CS5	Q96CS5 homo sapien
991	24	77.4	373	17	Q8U163	Q8U163 pyrococcus
992	24	77.4	375	5	Q21903	Q21903 caenorhabd
993	24	77.4	375	10	Q91TF5	Q91TF5 arabidopsis
994	24	77.4	375	17	Q8RT20	Q8RT20 pyrococcus
995	24	77.4	376	4	Q96E06	Q96E06 homo sapien
996	24	77.4	377	16	Q8YFE2	Q8YFE2 anabaena sp
997	24	77.4	378	5	Q9NDU9	Q9NDU9 caenorhabd
998	24	77.4	379	2	Q48738	Q48738 lactococcus
999	24	77.4	380	16	Q9CK63	Q9CK63 pasteurella
1000	24	77.4	380	16	Q8R869	Q8R869 thermoaer

ALIGNMENTS

RESULT 1

Q9DOK5	PRELIMINARY;	PRT;	114 AA.
AC	Q9DOK5;		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	2610008E20R1X protein.		
GN	CHST1 OR 2610008E20R1K.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;		
RA	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shindagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Akawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Alakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,		
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
RA	Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyon P., Marchionni L., Mashima U., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		

RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK013444; BAB27556.1;
 DR MGD: MGI:1924219; Chst1.
 SQ SEQUENCE 114 AA; 13232 MW; ABA829F593B80892 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 114;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYEDL 6
 DB 3 VYEDL 8

Q9VWC2	PRELIMINARY;	PRT;	183 AA.
AC	Q9VWC2;		
DT	01-MAY-2000 (TREMBlrel. 13, Created)		
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)		
DT	01-MAY-2000 (TREMBlrel. 13, Last annotation update)		
DE	G13765 protein.		
GN	C613765.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephyroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RA	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne C.R., Miklos G.L.G.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKlos G.L.G.,		
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,		
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,		
RA	Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reichert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Shue E., Spirdling A.C., Stapleton M., Strong R., Sun E.,		
RA	Switskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,		
RT	"The genome sequence of Drosophila melanogaster."		
RL	Science 287:2185-2195(2000).		

DR EMBL: AE003614; AAF52399.1; -
DR FlyBase: FBgn0031828; CG13765.
SQ SEQUENCE 183 AA; 21351 MW; C353243BFAFEED8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 75 VREEDL 80

RESULT 3

ID 0980N6 PRELIMINARY; PRT; 281 AA.

DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SS00254.
GN SS00254.

OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=1142726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

DR EMBL: AE006661; AAK40593.1; -
DR InterPro: IPR002747; DUF62.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF01887; DUF62; 1.
DR ProDom: PD013834; DUF62; 1.
DR PROSITE: PS00307; LECTIN_LEGUM_BETA; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 281 AA; 31911 MW; 23B2433BE3F375F CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 17; Length 281;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 160 VREEDL 165

ID 011316 PRELIMINARY; PRT; 331 AA.

DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE B1-13 protein (Fragment).

OS Moluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; GSDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Moluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]

RP SEQUENCE FROM N.A.
MORELLA M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,

RA Lopez-Estebarez J.L., Esteban M., Martin-Gallardo A;
RT "A Random DNA Sequencing, Computer-based Approach for the Generation
RT of a Gene Map of Moluscum Contagiosum Virus.";
RL Virus Genes 0:0-0(1997).

DR EMBL: U86899; AAB57946.1; -
FT NON_TER 1 1
FT NON_TER 331 331
SQ SEQUENCE 331 AA; 37009 MW; E776B399590B997E CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 331;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 225 VREEDL 230

RESULT 5

ID 09Y5R3 PRELIMINARY; PRT; 386 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand
DE sulfotransferase GSF-3).
GN GSF3.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RX MEDLINE=99264336; PubMed=10330415;
RA Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R.,
RA Huang C.C., Kanagaki R., Rosen S.D., Hemmerich S.;
RT "Sulfotransferases of two specificities function in the reconstitution
RT of high endothelial cell ligands for L-selectin.";
RL J. Cell Biol. 145:899-910(1999).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
RA Rosen S.D.;
RT "Chromosomal Localization and Genomic Organization for the
RT Galactose/N-Acetylglucosamine/N-Acetylglucosamine 6-O-
RT Sulfotransferase Gene Family.";
RL Glycobiology 0:0-0(2001).

[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21332592; PubMed=11439191;
RA Yeh J.-C., Hiraoaka N., Petryniak B., Nakayama J., Ellis L.G.,
RA Rabuka D., Hindsgrau O., Marth J.D., Lowe J.B., Fukuda M.;
RT "Novel sulfated lymphocyte homing receptors and their control by a
RT core extension beta1,3-N-acetylglucosaminyltransferase.";
RL Cell 105:957-969(2001).

DR EMBL: AF131235; AAD33015.1; -
DR EMBL: AF280088; AAG48246.1; -
DR EMBL: AF149783; AAK48417.1; -
DR InterPro: IPR001092; HLH_Basic.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
KW Lectin; Selection; Transferase.

SQ SEQUENCE 386 AA; 45133 MW; 0C3B4022417143A CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 274 VREEDL 279

```

RESULT 6
Q90UP4 PRELIMINARY; PRT; 395 AA.
AC Q90UP4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase.
GN CHST5 OR I-GLCNAc-6-ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF176841; AAD56003.1; -.
DR EMBL; AF176840; AAD56002.1; -.
DR MGD; MGI:1931825; Chst5.
KW Transferase.
SQ SEQUENCE 395 AA; 44537 MW; 3FDF71EA3ED383BE CRC64;

Query Match 100.0%; Score 31; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VREEDL 6
Db 271 VREEDL 276

RESULT 7
O43916 PRELIMINARY; PRT; 411 AA.
AC O43916;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN CHST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97189336; PubMed=9117134;
RA Williams K.J.;
RT "Atherosclerosis: cell biology and lipoproteins.";
RL Curr. Opin. Lipidol. 7:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX Peng T., Tadas I., Williams K.J.;
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98070405; PubMed=9405439;
RA Fukuta M., Inazawa J., Torii T., Tsuzuki K., Shimada E., Habuchi O.;
RT "Molecular cloning and characterization of human keratan sulfate
Gal-6-sulfotransferase.";
RL J. Biol. Chem. 272:32321-32328(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=DUBILICAL VEIN ENDOTHELIUM;
MEDLINE=99168906; PubMed=10049591;

```

```

RA Li X., Tedder T.F.;
RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
endothelial cells: cDNA cloning, expression, and chromosomal
localization.";
RL Genomics 55:345-347(1999).
DR EMBL; U65637; AAC28776.1; -.
DR EMBL; AB003791; BAA24840.1; -.
DR EMBL; AF090137; AAD19878.1; -.
KW Transferase.
SQ SEQUENCE 411 AA; 46714 MW; F07D0A23B6338A09 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VREEDL 6
Db 300 VREEDL 305

RESULT 8
Q9E0C0 PRELIMINARY; PRT; 411 AA.
AC Q9E0C0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Galactose 6-O-sulfotransferase GST-1.
GN CHST1 OR GST1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=MAMMARY GLAND.
RA Hemmerich S., Lee J.-K., Bhakta S., Bistup A., Ruddle N.R.,
Rosen S.D.;
RT "Chromosomal localization and genomic organization of the galactose/N-
acetylglucosamine/N-acetylglucosamine 6-O-sulfotransferase gene
family.";
RL Glycobiology 11:75-87(2000).
DR EMBL; AF280087; AAC48245.1; -.
DR MGD; MGI:1924219; Chst1.
KW Transferase.
SQ SEQUENCE 411 AA; 46903 MW; B1AE590EF5B9C6DC CRC64;

Query Match 100.0%; Score 31; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VREEDL 6
Db 300 VREEDL 305

RESULT 9
Q9RV03 PRELIMINARY; PRT; 428 AA.
AC Q9RV03;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein DR0969.
GN DR0969.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
MEDLINE=20036896; PubMed=10567266;

```

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vasmatachean J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AF001949; AAF10548.1; -.
 DR TIGR: DR0969; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 428 AA; 46258 MW; 00ED2F1682C0DF45 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 120 VREEDL 125

RESULT 10

ID 09UED5 PRELIMINARY; PRT; 484 AA.
 AC 09UED5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST).
 GN GN6ST.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=9391845; PubMed=9722682;
 RA Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q.,
 RA Mitsuoka C., Kanagaki R., Habuchi O., Yokoyama I., Yamamura K.,
 RA Ozaki T., Nakagawara A., Kadomatsu K., Muramatsu T.;
 RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the
 RT biosynthesis of 6-sulfo sialyl Lewis X: Molecular cloning, chromosomal
 RT mapping, and expression in various organs and tumor cells.";
 RL J. Biochem. 124:670-678(1998).
 DR EMBL: AB014679; BAA34265.1; -.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 484 AA; 53641 MW; D077EBBD7645F2F0 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 484;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 377 VREEDL 382

RESULT 11

ID 09JTN9 PRELIMINARY; PRT; 486 AA.
 AC 09JTN9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE LD3505.P.
 GN CG9551.
 OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brickstein P., Hong L., Appayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Fise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Paclio J., Paragas V., Park S., Photanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisier S.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY058647; AAL13876.1; -.
 DR FlyBase: FBgn0031827; CG9551.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 486 AA; 53539 MW; BEDCC96F096746EB CRC64;

Query Match 100.0%; Score 31; DB 5; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 378 VREEDL 383

RESULT 12

ID 09Y6F2 PRELIMINARY; PRT; 530 AA.
 AC 09Y6F2; Q9GZNS;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Carbohydrate sulfotransferase 2 (N-acetylglucosamine
 DE 6-O-sulfotransferase).
 GN CHST2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE=99168906; PubMed=10049591;
 RA Li X., Tedder T.F.;
 RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
 RT endothelial cells: cDNA cloning, expression, and chromosomal
 RT localization.";
 RL Genomics 55:345-347(1999).
 RN [2]
 RP SEQUENCE OF 48-530 FROM N.A.
 RC TISSUE-PLACENTA;
 RX MEDLINE=20498786; PubMed=11042394;
 RA Sakaguchi H., Kitagawa H., Sugahara K.;
 RT "Functional expression and genomic structure of human N-
 RT acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-
 RT acetylglucosamine at the nonreducing end of an N-acetylglucosamine
 RT sequence.";
 RL Biochim. Biophys. Acta 1523:269-276(2000).
 DR EMBL: AF083066; AAD20981.1; -.
 DR EMBL: AB021125; BAB16887.1; -.
 DR EMBL: AB021124; BAB16886.1; -.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 530 AA; 57857 MW; A82CA2279D56551B CRC64;

Query Match 100.0%; Score 31; DB 4; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 423 VREEDL 428

RESULT 13

ID 088276 PRELIMINARY; PRT; 530 AA.

AC 088276;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE N-acetylglucosamine-6-O-sulfotransferase long form.
GN CHST2.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=WHOLE EMBRYOS;
RX MEDLINE=98380482; PubMed=9712885;

RA Uchimura K., Muramatsu H., Kadomatsu K., Fan Q., Kurosawa N.,
RA Mitsuoka C., Kannagi R., Habuchi O., Muramatsu T.,
RT "Molecular cloning and characterization of an N-acetylglucosamine-6-O-

RT sulfotransferase.";
RL J. Biol. Chem. 273:22577-22583(1998).

DR EMBL; AB011452; BAA52138.1;
DR EMBL; AB011452; BAA52139.1;
DR MGI; MGI:1891160; Chst2.

DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.

KM TRANSFERASE 530 AA; 57814 MW; A113E1B735C363EC CRC64;
SQ SEQUENCE

Query Match 100.0%; Score 31; DB 11; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 423 VREEDL 428

RESULT 14

ID 09Y4C5 PRELIMINARY; PRT; 531 AA.

AC 09Y4C5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Long form of N-acetylglucosamine-6-O-sulfotransferase
DE (G1CNA6ST).

GN G1CNA6ST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;
RX MEDLINE=98391845; PubMed=9722682;

RA Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q.,
RA Mitsuoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K.,
RA Ozaki T., Nakagawara A., Kadomatsu K., Muramatsu T.,
RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the
RT biosynthesis of 6-sulfo sialyl Lewis x: Molecular cloning, chromosomal
RT mapping, and expression in various organs and tumor cells.";
RL J. Biochem. 124:670-678(1998).

DR EMBL; AB014680; BAA34266.1;
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.

KW Transferase.
SQ SEQUENCE 531 AA; 58768 MW; D2EDB74E95B5162F CRC64;

Query Match 100.0%; Score 31; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 424 VREEDL 429

RESULT 15

ID 09Y158 PRELIMINARY; PRT; 597 AA.

AC 09Y158;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE BCDNA:GH02901.
GN BCDNA:GH02901 OR CG9009.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazek R.G., Butenoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Mostreli M., Pacle J.M.,
RA Park S., Segueira A., Sethi H., Snir E., Svirsks R.R., Weinburg T.,
RA Ceilinker S.E.;
RT "Full length Drosophila melanogaster cDNA sequence.";
RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF145610; AAD38585.1;
DR HSSP: P08659; ILCT.
DR FlyBase: FBgn0027601; BCDNA:GH02901.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00301; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.

SQ SEQUENCE 597 AA; 65570 MW; 088010B38190DB22 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 229 VREEDL 234

RESULT 16

ID 025212 PRELIMINARY; PRT; 631 AA.

AC 025212;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein HP0465.

GN HP0465.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;

RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Kap P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 368:539-547(1997).
 DR EMBL: AE000561; AAD07527.1; -.
 DR TIGR: HP0465; -.
 DR InterPro: IPR000834; Zn_carboxept.
 DR ProSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 631 AA; 72171 MW; 81380B27F060870B CRC64;

Query Match 100.0%; Score 31; DB 16; Length 631;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
 DB 182 VREYDL 187

RESULT 17
 O9Y884 PRELIMINARY; PRT; 649 AA.
 AC O9V884;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG18635 protein.
 GN CG18635.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Adiyanti A., An H.-T., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modyarty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switksis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Wortley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003802; AAF57791.1; -.
 DR FlyBase: FBgn034279; CG18635.
 SQ SEQUENCE 649 AA; 74414 MW; 6471AC52B499A9E9 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 649;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
 DB 332 VREYDL 337

RESULT 18
 O98200 PRELIMINARY; PRT; 748 AA.
 AC O98200;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE MC032L.
 GN MC032L.
 OS Moluscum contagiosum virus subtype 1 (MCV1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Molluscipoxvirus.
 OC NCBI_TaxId=10280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96325459; PubMed=8670425;
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
 RA Moss B.;
 RT "Genome sequence of a human tumorigenic poxvirus: Prediction of
 RT specific host response-evasion genes.";
 RL Science 273:813-816(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
 RA Moss B.;
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U60315; AAC55160.1; -.
 SQ SEQUENCE 748 AA; 84411 MW; 61A2741F62C0A5BA CRC64;

Query Match 100.0%; Score 31; DB 12; Length 748;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYDL 6
 DB 617 VREYDL 622

RESULT 19
 O45412 PRELIMINARY; PRT; 973 AA.
 ID O45412;
 AC O45412;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE F28B1.2 protein.
 GN F28B1.2.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Rhabditinae; Rhabditidae;
 OC NCBI_TaxId=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81517; CAB04209.1; -
 SO SEQUENCE 973 AA; 109292 MW; 74F2C8C63BA94067 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 973;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VRYEDL 6
 Db 828 VRYEDL 833

RESULT 20
 OS4299 PRELIMINARY; PRT; 6260 AA.
 ID OS4299;
 AC OS4299;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Polypeptide synthase.
 GN RABC.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NX NCBI_TaxID=1912;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 5491;
 RX MEDLINE=9537374; PubMed=7644502;
 RA Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
 Haydock S.F., O'Leary M., Caffrey P., Cortes J., Lester J.B.,
 Boehm G.A., Staunton J., Leadlay P.F.,
 "The biosynthetic gene cluster for the polypeptide immunosuppressant
 RT rapamycin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 5491;
 RX MEDLINE=96186895; PubMed=8635730;
 RA Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
 Koenig A., Staunton J., Leadlay P.F.,
 "Organisation of the biosynthetic gene cluster for rapamycin in
 RT Streptomyces hygroscopicus: analysis of genes flanking the polypeptide
 RT synthase.";
 RL Gene 169:1-7(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 5491;
 RX MEDLINE=96186896; PubMed=8635756;
 RA Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
 Khaw L.E., Staunton J., Leadlay P.F., Staunton J., Leadlay P.F.,
 "Organization of the biosynthetic gene cluster for rapamycin in
 RT Streptomyces hygroscopicus: analysis of the enzymatic domains in the
 RT modular polypeptide synthase.";
 RL Gene 169:9-16(1996).
 DR EMBL: X86780; CA60462.1; -
 DR HSSP; P25715; IMIA.
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR002085; adh_zn_family.
 DR InterPro: IPR004410; Pabd.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl-synt.

DR InterPro: IPR003880; Ppantn_attach.
 DR InterPro: IPR002364; QOR_zeta_crystal.
 DR InterPro: IPR000634; S/T_dehydratase.
 DR Pfam: PF00698; Acyl_transf. 4.
 DR Pfam: PF00107; adh_zinc. 1.
 DR Pfam: PF00109; ketoacyl-synt. 4.
 DR Pfam: PF02801; ketoacyl-synt_C. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR TIGRfam: TIGR00128; Iabd. 4.
 DR PROSITE: PS00075; ACP_DOMAIN. 4.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE. 4.
 DR PROSITE: PS00165; DEHYDRATASE_SER_THR; UNKNOWN.1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
 DR PROSITE: PS00012; PHOSPHOPANTHETINE. 4.
 DR PROSITE: PS01162; QOR_ZETA_CRYSTAL. 1.
 DR Phosphopantetheine transferase.
 SW SEQUENCE 6260 AA; 659509 MW; D76C002F91F2080 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 6260;
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VRYEDL 6
 Db 4365 VRYEDL 4370

RESULT 21
 OS4299 PRELIMINARY; PRT; 194 AA.
 ID OS4299;
 AC OS4299;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein PAE3336.
 GN PAE3336.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 NX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladder H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.,
 "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*
 RT aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AE009925; AAL6489.1;
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 194 AA; 22556 MW; C3B5B464EEES58D0 CRC64;

Query Match 96.8%; Score 30; DB 17; Length 194;
 Best Local Similarity 83.3%; Pred. No. 72;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VRYEDL 6
 Db 83 VRYEDL 88

RESULT 22
 OS4299 PRELIMINARY; PRT; 199 AA.
 ID OS4299;
 AC OS4299;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 220002K05R1K protein.
 GN 220002K05R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia: Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 Saito T., Okazaki Y., Gajobori T., Bono H., Kasukawa T., Saito R.,
 Radtka K., Matsuda H.A., Ashbunner M., Balalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., Kang B., Kochia H.,
 Kuenl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,
 Schriml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aoto H., Balarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hall D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK008621; BAB25783.1; -;
 DR MGD: MGI:1916387; 2200002K05Rik.
 DR InterPro: IPR003599; I9.
 DR SMART: SM00409; IG; 1.
 SQ SEQUENCE 199 AA; 22564 MW; 77400D2FF9AD4642 CRC64;
 Query Match 96.8%; Score 30; DB 11; Length 199;
 Best Local Similarity 83.3%; Pred. No. 74;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VREYEDL 6
 Db 140 IREYEDL 145
 ID :|||||
 AC 09VMC4 PRELIMINARY; PRT; 363 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG9550 protein.
 GN CG9550.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Adams M.D., Celinker S.E., Holt R.A., Hoskins R.A., Galle R.F.,
 Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers J.H.C., Blake J., Champ E., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Doudson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003614; AAF52397.1; -;
 DR FLYBase: FBgn0031826; CG9550.
 SQ SEQUENCE 363 AA; 41932 MW; C3CC09E608FEF8AC CRC64;
 Query Match 96.8%; Score 30; DB 5; Length 363;
 Best Local Similarity 83.3%; Pred. No. 1,4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VREYEDL 6
 Db 258 IREYEDL 263
 ID :|||||
 AC 08TUC3 PRELIMINARY; PRT; 540 AA.
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Methylamine methyltransferase corrinoid activation protein.
 GN RMA OR MA0150.
 GN RMA OR MA0150.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11937238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
 Fitzhugh W., Calvo S., Engels R., Smirnov S., Altouf D., Brown A.,
 Allen N., Naylor J., Stange-Thomann N., DeArrello K., Johnson R.,
 Linton L., McEwan P., McKernan K., Talmas J., Tirrell A., Ye W.,
 Zimmer A., Barber R.D., Cann I., Graham D.E., Girhane D.A., Guss A.M.,
 Heiderich J.R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 McCall W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL: AB010672; AA003603.1; -;
 DR FLYBase: FBgn0031826; CG9550.
 SQ SEQUENCE 540 AA; 58864 MW; 7B3D323035D5227B CRC64;
 Query Match 96.8%; Score 30; DB 17; Length 540;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 326 IRYEDL 331

RESULT 25

09A776 PRELIMINARY; PRT; 579 AA.

AC 09A776;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Oxidoreductase, GMC family.
GN CC1634.

OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;

RN 11

SEQUENCE FROM N.A.

RC STRAIN=ANCC 19089 / CB15;
MEDLINE=21173698; PubMed=11259647;

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Ptoczka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uitterback T., Tran K., Wolf A., Vamathevan J., Emiolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR EMBL; AE005839; AAK23612.1; -

DR TIGR; CC1634; -

DR InterPro; IPR000205; NAD_binding.

KW Complete proteome.

QY SEQUENCE 579 AA; 64397 MW; 7125C4DPAD618F10 CRC64;

Query Match
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 153 IRYEDL 158

RESULT 26

08SRF6 PRELIMINARY; PRT; 959 AA.

AC 08SRF6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SEC31-like protein involved in vesicular transport from ER to
DE golgi.
GN ECU08.0210.

OS Eucephalitozoon cuniculi.

OC Eukaryota; Microsporidia; Unikaryoniidae; Eucephalitozoon.

OX NCBI_TaxID=6035;

RN 11

SEQUENCE FROM N.A.

RC STRAIN=GB-M1.

RA Genoscope; (APR-2001) to the EMBL/GenBank/DBJ databases.

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN 12

SEQUENCE FROM N.A.

RC STRAIN=GB-M1.

RA MEDLINE=21576510; PubMed=11719806;

RA Kainka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,

RA Prenster G., Barbe V., Peyretailade E., Brottier P., Wincker P.,

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.,
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Eucephalitozoon cuniculi."
RL Nature 414:450-453(2001).
DR EMBL; AL590448; CAD26327.1; -
SQ SEQUENCE 959 AA; 104621 MW; 6EC8782742424322 CRC64;

Query Match
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 584 IRYEDL 589

RESULT 27

023559 PRELIMINARY; PRT; 1322 AA.

AC 023559;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical 150.2 kDa protein.
GN AT4G17140.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN 11

SEQUENCE FROM N.A.

RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,

RA Kreis M., Kavanagh T., Entlan K.D., Rieger M., James R.,

RA Pridmorenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,

RA Jones J., Palmer K., Ansorge W., Delseny M., Bancroft I., Meves H.W.,

RA Scheller C., Chaiwatzis N.,

RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN 12

SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z97342; CAB10495.1; -

DR EMBL; AL161545; CAB80985.1; -

DR InterPro; IPR01849; PH.

DR Pfam; PF00169; PH.1.

DR SMART; SM00233; PH.1.

DR PROSITE; PS50003; PH_DOMAIN; 1.

KW Hypothetical protein.

SQ SEQUENCE 1322 AA; 150221 MW; F70DE47DB21CEAF4 CRC64;

Query Match
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 21 IRYEDL 26

RESULT 28

09A994 PRELIMINARY; PRT; 234 AA.

AC 09A994;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein CC1100.

GN CC1100.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.

OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA MEDLINE=2136508; PubMed=11481431;
 RX Fitan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhieslter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
 RA Golding B., Puchler A.;
 RT "The complete sequence of the 1.683-kb pSYM megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
 DR EMBL: AF603647; CAC49830.1;
 DR InterPro: IPR000567; SBP_bac_1;
 DR Pfam: PF01547; SBP_bacterial_1;
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 39502 MW; B5B32E6094D7AFCB CRC64;

Query Match 93.5%; Score 29; DB 16; Length 355;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 63 VREEDM 68

RESULT 32
 ID 081244 PRELIMINARY; PRT; 393 AA.
 AC 081244;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Polygalacturonase precursor.
 GN Mpg1.
 OS Cucumis melo (Muskmelon).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hadfield K.A., Rose J.K.C., Yaver D.S., Berka R.M., Bennett A.B.;
 RT "Polygalacturonase gene expression in ripe melon fruit supports a role
 RT for polygalacturonase in ripening-associated pectin disassembly."
 RL Plant Physiol. 116:0-0(1998)
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL: AF062465; AAC26510.1;
 DR InterPro: IPR000743; GH28.
 DR InterPro: IPR000408; Reg_chr_condens.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 DR PROSITE: PS00626; RCL1_2; UNKNOWN_1.
 KW Cell wall; Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1
 FT CHAIN 1 393 POTENTIAL.
 SQ SEQUENCE 393 AA; 42145 MW; B94F5635D5EB8AC7 CRC64;

Query Match 93.5%; Score 29; DB 10; Length 393;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 329 VREEDI 334

RESULT 33
 ID 022818 PRELIMINARY; PRT; 405 AA.

AC 022818;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Putative polygalacturonase.
 GN AT2G43860.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rongling C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Talion L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL: AC002333; AAB64020.1;
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 KW Cell wall; Glycosidase; Hydrolase
 SQ SEQUENCE 405 AA; 43462 MW; 556D6557044D195 CRC64;

Query Match 93.5%; Score 29; DB 10; Length 405;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
 DB 335 VREEDI 340

RESULT 34
 ID P96458 PRELIMINARY; PRT; 154 AA.
 AC P96458;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE HSPR.
 GN HSPR.
 OS Streptomyces albus G.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97157943; PubMed=9004222;
 RA Grandvalet C., Servant P., Mazodier P.;
 RT "Distruption of hspR, the repressor gene of the dnaX operon in
 RT Streptomyces albus G.";
 RL Mol. Microbiol. 23:77-84 (1997).
 CC -1- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: U43299; AAB48078.1;
 DR InterPro: IPR000551; HTH_Merr.
 DR Pfam: PF00376; merr; 1.
 DR SMART: SM00422; HTH_MERR; 1.

DR PROSITE: PS00552: HTH_MERR_FAMILY: 1.
 KM DNA-binding: Transcription regulation.
 SQ SEQUENCE 154 AA; 17400 MW; F5C64F168BCF30B8 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 2; Length 154;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 :|||||
 DB 132 VRYEDV 137

RESULT 35
 ID 028741 PRELIMINARY; PRT; 195 AA.
 AC 028741:
 DT 01-JUN-1998 (TREMBLrel. 05, Created)
 DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein AF1531.
 GN AF1531.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.R., Peterson J.D.,
 RA Richardson D.D., Kellavag A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirsnes E.F., Dougherty B.A., McInerney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kalne B.P., Sykes S.W.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE000997; AAB89717.1; -
 DK TIGR: AF1531; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 195 AA; 23194 MW; C244F95420565E2C CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 17; Length 195;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 :|||||
 DB 87 LRYEDL 92

RESULT 36
 ID 0976N0 PRELIMINARY; PRT; 228 AA.
 AC 0976N0:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein ST0160.
 GN ST0160.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / ?;
 RA PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takehashi M.,
 RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermocidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000981; BAB55116.1; -
 DR InterPro: IPR000209; Peptidase_S8.
 DR PROSITE: PS00136; SUBTILASE_ASF; UNKNOWN_1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 228 AA; 27361 MW; 7BCE2DC618163CE6 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 17; Length 228;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 :|||||
 DB 212 VRYEDL 217

RESULT 37
 ID 09CEP1 PRELIMINARY; PRT; 233 AA.
 AC 09CEP1:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Regulatory protein.
 GN YSPD OR L11795.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bojorin A., Wincker P., Manger S., Jallion O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis TL1403.";
 RL Genome Res. 11:731-753(2001).
 DK EMBL: AE006409; AAK05893.1; -
 KW Complete proteome.
 SQ SEQUENCE 233 AA; 27077 MW; 483DAD4E8896A8BD CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 16; Length 233;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 :|||||
 DB 146 IRYEDI 151

RESULT 38
 ID 09RC18 PRELIMINARY; PRT; 273 AA.
 AC 09RC18:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Thermostable chitosanase.
 GN CHOK.
 OS Bacillus sp. KB-CO4.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.

OC NCBI_TaxID=111825;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KFB-CO4;
 RX MEDLINE=20422179; PubMed=10966383;
 RA Yoon H.G., Kim H.Y., Lim Y.H., Kim H.K., Shin D.H., Hong B.S.,
 RA Cho H.Y.,
 RT "Thermotable chitosanase from *Bacillus* sp. Strain CK4: Cloning and
 RT expression of the gene and characterization of the enzyme.";
 RL Appl. Environ. Microbiol. 66:3727-3734(2000).
 DR EMBL: AF160195; AAF24188.1;
 DR InterPro: IPR000400; GH_46.
 DR Pfam: PF01374; Glyco_hydro_46; 1.
 DR PROSITE: P560000; CHITOSANASE_46.80; UNKNOWN_1.
 SQ SEQUENCE 273 AA; 30552 MW; C939168C9DD27B9 CRC64;

Query Match 90.3%; Score 28; DB 2; Length 273;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 218 VRYDDL 223

RESULT 39
 O9YTO1 PRELIMINARY; PRT; 273 AA.

AC O9YTO1;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Mitogen.
 OS Ateles herpesvirus 3.
 OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=85618;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=73;
 RX MEDLINE=20091363; PubMed=10623770;
 RA Albrecht J.C.;
 RT "Primary structure of the Herpesvirus Ateles genome.";
 RL J. Virol. 74:1033-1037(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=73;
 RA Albrecht J.C., Fleckenstein B.;
 RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083424; AAC95536.1;
 DR InterPro: IPR001213; MMTV_SAg.
 DR Pfam: PF01054; MMTV_SAg; 1.
 SQ SEQUENCE 273 AA; 30599 MW; E0C418C2DEA56F1A CRC64;

Query Match 90.3%; Score 28; DB 2; Length 273;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 208 VRYDDL 213

RESULT 40
 O9ET84 PRELIMINARY; PRT; 278 AA.

AC O9ET84;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Chitosanase precursor.
 OS *Bacillus amyloliquefaciens*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae: *Bacillus*.
 OX NCBI_TaxID=1390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seki K., Kuriyama H., Uchida Y.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Seki K., Kuriyama H., Okuda T., Uchida Y.;
 RT "Molecular Cloning of the Gene Encoding Chitosanase from *Bacillus*
 RT *amyloliquefaciens* UTK.";
 RL (in) Unknown A. (eds.);
 RL Advances in Chitin Science 2, pp.284-289, Unknown Publisher (1997).
 DR EMBL: AB051574; BAB19276.1;
 DR HSSP: P33665; ICHR.
 DR InterPro: IPR000400; GH_46.
 DR Pfam: PF01374; Glyco_hydro_46; 1.
 DR PROSITE: P560000; CHITOSANASE_46.80; UNKNOWN_1.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 278 AA; 31404 MW; A742779DB91B51B CRC64;

Query Match 90.3%; Score 28; DB 2; Length 278;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 223 VRYDDL 228

RESULT 41
 O9I732 PRELIMINARY; PRT; 306 AA.

AC O9I732;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE Hypothetical protein PA0100.
 GN PA0100.
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01.
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stever C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004448; AAG03490.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 306 AA; 33723 MW; E9EF587A69005111 CRC64;

Query Match 90.3%; Score 28; DB 16; Length 306;
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 16 VRYDDL 21

RESULT 42
 O82330 PRELIMINARY; PRT; 333 AA.


```
AC 082330;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative steroid sulfotransferase.
GN AT2G14920.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.F., Feldlyum T.V.,
RA Beil C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005396; AAC61289.1; -
DR HSSP; P50224; ICDM.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferrase.
SQ SEQUENCE 333 AA; 38924 MW; B13B671253C619CF CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 10; Length 333;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 232 MREYDL 237

RESULT 43
ID 045982 PRELIMINARY; PRT; 346 AA.
AC 045982;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical 40.0 kDa protein.
GN ZC513.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Wu X., Le T.T.;
RT "The sequence of C. elegans cosmid ZC513.";
```

```
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53155; AAC48271.1; -
DR InterPro; IPR003002; 7TM_chemol.
DR InterPro; IPR000168; 7TM_nematode.
DR Pfam; PF01461; 7tm_4; 1.
KW Hypothetical protein.
SQ SEQUENCE 346 AA; 40010 MW; 30A80441AA784868 CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 5; Length 346;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 79 LREYDL 84

RESULT 44
ID 09FL45 PRELIMINARY; PRT; 354 AA.
AC 09FL45;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Steroid sulfotransferase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010697; BAB11159.1; -
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferrase.
SQ SEQUENCE 354 AA; 40887 MW; E540A1449966A36C CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 10; Length 354;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 249 LREYDL 254

RESULT 45
ID 0940P4 PRELIMINARY; PRT; 359 AA.
AC 0940P4;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Ate907010/MO9.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

OC eutrosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banb J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kaniya A.,
 RA Kallin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinnokki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis cDNA clones.",
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY054219; AL06879.1; -
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransferase.1.
 DR ProDom: PD001218; Sulfotransferase.1.
 SO SEQUENCE 359 AA; 41378 MW; 0614AF7987C3F4A4 CRC64;
 Query Match 90.3%; Score 28; DB 10; Length 359;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 Db 254 LRYEDL 259
 RESULT 46
 O9R111 PRELIMINARY; PRT; 388 AA.
 AC O9R111
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE L-selectin ligand sulfotransferase.
 GN CHST4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9361934; PubMed=10435581;
 RA Hiraoka N., Petryniak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,
 RA Itawa D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.,
 RT "A novel, high endothelial venule-specific sulfotransferase expresses
 RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.",
 RL Immunity 11:79-89(1999).
 DR EMBL: AF109155; MAD45579.1; -
 DR MGD: MGI:1349479; Chst4.
 KW Lectin; Selectin; Transferase.
 SO SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884ADE CRC64;
 Query Match 90.3%; Score 28; DB 11; Length 388;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 Db 273 LRYEDL 278
 RESULT 47
 O9WUE5 PRELIMINARY; PRT; 388 AA.
 AC O9WUE5
 DT 01-NOV-1998 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE N-acetylglucosamine 6-O-sulfotransferase (Carbohydrate Chondroitin
 DE 6/eratan) sulfotransferase 4).

GN CHST4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6; TISSUE=TONSTL;
 RC MEDLINE=99264336; PubMed=10330415.
 RA Bistrup A., Bhakta S., Lee J.-K., Belov Y.Y., Gunn M.D., Zuo F.-R.,
 RA Huang C.-C., Kannagi R., Rosen S.D., Hemmerich S.,
 RT "Sulfotransferases of two specificities function in the reconstitution
 RT of high endothelial cell ligands for L-selectin.",
 RL J. Cell Biol. 145:899-910(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner I., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 DR EMBL: AF131236; MAD33016.1; -
 DR EMBL: AK009113; BAB26078.1; -
 DR MGD: MGI:1349479; Chst4.
 KW Transferase.
 SO SEQUENCE 388 AA; 44694 MW; D1E9D796DF8574D CRC64;
 Query Match 90.3%; Score 28; DB 11; Length 388;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 Db 273 LRYEDL 278
 RESULT 48
 O97B92 PRELIMINARY; PRT; 391 AA.
 AC O97B92
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Molybdenum cofactor biosynthesis protein moea.
 GN TYO565 OR TYG0553325.
 OS Thermoplasma volcanium.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=50339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;
 RX MEDLINE=20570466; PubMed=11121031;
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
 RA Kawashima-Onya Y., Matsunabe K., Yamazaki M., Kanehori K., Kawamoto T.,
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.,

RT "Archaeal adaptation to higher temperatures revealed by genomic
 RT sequence of Thermoplasma volcanium";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
 DR EMBL: AF000992; BAB59707.1; -;
 DR InterPro: IPR001453; MOCF_biosynth.
 DR InterPro: IPR005111; Moea_C.
 DR InterPro: IPR005110; Moea_N.
 DR Pfam: PF00994; MOCF_biosynth. 1.
 DR Pfam: PF03454; Moea_C; 1.
 DR Pfam: PF03453; Moea_N; 1.
 DR Prodom: PD002460; MOCF_biosynth. 1.
 DR Trigram: TRIGR00177; molyb_syn; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 42773 MW; 79BC9BA181FE99B CRC64;

QY 1 VREEDL 6
 :|||||:
 DB 56 VREEDV 61

Query Match 90.3%; Score 28; DB 17; Length 391;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 49
 O9V129 PRELIMINARY; PRT; 458 AA.
 ID O9V129;
 AC O9V129;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Tlud related protein.
 GN PAB1955.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248284; CAB49522.1; -;
 DR InterPro: IPR002510; PmbA_Tlud.
 DR Pfam: PF01523; PmbA_Tlud; 1.
 KW Complete proteome.
 SQ SEQUENCE 458 AA; 51567 MW; 3B520B59A370DB4D CRC64;

Query Match 90.3%; Score 28; DB 17; Length 458;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 :|||||:
 DB 25 LREEDL 30

RESULT 50
 O93N59 PRELIMINARY; PRT; 472 AA.
 ID O93N59;
 AC O93N59;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 55.0 kDa protein.
 OS Coxiella burnetii.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Coxiella group; Coxiella.
 OX NCBI_TaxID=777;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
 RT "A chromosomal DNA deletion explains the phenotype of the Coxiella
 RT burnetii phase II variant.";
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF387640; AAK71263.1; -;
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 472 AA; 54997 MW; 6E1507D77A53FA03 CRC64;

QY 1 VREEDL 6
 :|||||:
 DB 288 IREED 293

Query Match 90.3%; Score 28; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 5.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Search completed: February 20, 2003, 13:33:48
 Job time : 32.5714 secs

This Page Blank (uspio)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:31 ; Search time 6.85714 Seconds

(without alignments)
25.745 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued Patents,AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	6	US-09-045-284A-9
2	31	100.0	4	US-09-045-284A-2
3	31	100.0	386	US-09-190-911-1
4	31	100.0	411	US-09-015-188-2
5	31	100.0	483	US-09-263-023-2
6	31	100.0	483	US-09-471-867-2
7	31	100.0	484	US-09-263-023-4
8	31	100.0	484	US-09-471-867-4
9	28	90.3	306	US-08-867-030B-13
10	28	90.3	306	US-08-867-030B-13
11	28	90.3	458	US-08-655-879-13
12	28	90.3	479	US-08-899-514-2
13	28	90.3	479	US-08-899-514-2
14	28	90.3	626	US-08-472-934-6
15	28	90.3	626	US-08-472-934-6
16	28	90.3	626	US-08-472-934-6
17	28	90.3	626	US-08-472-934-6
18	28	90.3	626	US-08-472-934-6
19	28	90.3	626	US-08-472-934-6
20	28	90.3	626	US-08-472-934-6
21	27	87.1	295	US-08-411-607A-2
22	27	87.1	308	US-08-927-433-9
23	27	87.1	308	US-08-927-433-9
24	27	87.1	308	US-08-927-433-9
25	27	87.1	308	US-08-927-433-9
26	27	87.1	308	US-08-927-433-9
27	87.1	606	4	US-09-041-236-4

28	27	87.1	983	4	US-09-412-554A-2	Sequence 2, Appli
29	27	87.1	1119	4	US-09-396-651B-2	Sequence 2, Appli
30	26	83.9	416	1	US-08-117-083-60	Sequence 60, Appli
31	26	83.9	465	2	US-08-588-983-20	Sequence 22, Appli
32	26	83.9	465	2	US-08-588-983-22	Sequence 22, Appli
33	26	83.9	465	2	US-08-588-976-20	Sequence 20, Appli
34	26	83.9	465	2	US-08-588-976-22	Sequence 22, Appli
35	26	83.9	919	2	US-08-588-993-9	Sequence 9, Appli
36	26	83.9	919	2	US-08-588-993-12	Sequence 12, Appli
37	26	83.9	919	2	US-08-588-976-9	Sequence 9, Appli
38	26	83.9	919	2	US-08-588-976-12	Sequence 12, Appli
39	26	83.9	1174	1	US-08-040-751-3	Sequence 3, Appli
40	26	83.9	1174	1	US-08-291-368-2	Sequence 2, Appli
41	26	83.9	1174	1	US-08-962-190-2	Sequence 2, Appli
42	26	83.9	1174	5	PCR-US95-10310-2	Sequence 2, Appli
43	26	83.9	1174	6	5164180-4	Patent No. 5164180
44	26	83.9	2548	4	US-09-172-422-1	Sequence 1, Appli
45	25	80.6	103	4	US-09-370-253-14	Sequence 14, Appli
46	25	80.6	171	4	US-09-352-990-20	Sequence 20, Appli
47	25	80.6	182	2	US-08-698-805-4	Sequence 4, Appli
48	25	80.6	248	4	US-09-210-7248A-2	Sequence 2, Appli
49	25	80.6	249	1	US-07-687-072B-2	Sequence 2, Appli
50	25	80.6	249	1	US-08-466-444-2	Sequence 2, Appli
51	25	80.6	322	4	US-09-134-001C-3194	Sequence 3194, Ap
52	25	80.6	391	4	US-09-134-001C-4880	Sequence 4880, Ap
53	25	80.6	400	4	US-09-352-990-18	Sequence 18, Appli
54	25	80.6	410	4	US-09-352-990-16	Sequence 16, Appli
55	25	80.6	432	4	US-09-370-253-12	Sequence 12, Appli
56	25	80.6	473	4	US-09-637-118B-2	Sequence 2, Appli
57	25	80.6	476	3	US-09-188-579-114	Sequence 114, App
58	25	80.6	476	4	US-09-315-444-114	Sequence 114, App
59	25	80.6	476	4	US-09-134-218-6	Sequence 6, Appli
60	25	80.6	476	4	US-09-721-362-114	Sequence 114, App
61	25	80.6	487	1	US-08-485-107-2	Sequence 2, Appli
62	25	80.6	608	4	US-09-134-001C-3855	Sequence 3855, Ap
63	25	80.6	615	1	US-08-484-105-6	Sequence 6, Appli
64	25	80.6	615	1	US-08-484-106-6	Sequence 6, Appli
65	25	80.6	853	4	US-09-254-352B-17	Sequence 17, Appli
66	25	80.6	854	4	US-09-254-352B-18	Sequence 18, Appli
67	25	80.6	883	1	US-08-106-433A-2	Sequence 2, Appli
68	25	80.6	883	4	US-09-254-352B-16	Sequence 16, Appli
69	25	80.6	1529	4	US-09-215-654-1	Sequence 1, Appli
70	25	80.6	2233	2	US-08-569-853-1	Sequence 1, Appli
71	25	80.6	2233	2	US-08-569-853-2	Sequence 2, Appli
72	25	80.6	2233	3	US-08-987-439-1	Sequence 1, Appli
73	24	77.4	20	1	US-08-171-299B-11	Sequence 11, Appli
74	24	77.4	26	4	US-08-858-207A-466	Sequence 466, App
75	24	77.4	52	4	US-09-369-494-17	Sequence 17, Appli
76	24	77.4	52	4	US-09-369-494-15	Sequence 15, Appli
77	24	77.4	93	4	US-08-936-165A-466	Sequence 466, App
78	24	77.4	94	4	US-08-055-006-21	Sequence 21, Appli
79	24	77.4	109	4	US-09-134-001C-5334	Sequence 5334, Ap
80	24	77.4	115	3	US-08-483-749A-6	Sequence 6, Appli
81	24	77.4	125	1	US-08-225-989-20	Sequence 20, Appli
82	24	77.4	125	1	US-08-580-923-20	Sequence 20, Appli
83	24	77.4	125	1	US-08-580-923-20	Sequence 20, Appli
84	24	77.4	125	4	US-09-078-785-20	Sequence 20, Appli
85	24	77.4	141	4	US-09-188-930-327	Sequence 327, App
86	24	77.4	147	6	5248606-46	Patent No. 5248606
87	24	77.4	148	3	US-08-584-031-12	Sequence 12, Appli
88	24	77.4	160	4	US-09-134-001C-4327	Sequence 4327, Ap
89	24	77.4	198	1	US-08-642-255-111	Sequence 111, App
90	24	77.4	198	1	US-08-397-633A-11	Sequence 41, Appli
91	24	77.4	198	1	US-08-397-633A-52	Sequence 52, Appli
92	24	77.4	203	4	US-09-134-001C-3696	Sequence 3696, Ap
93	24	77.4	215	1	US-08-225-989-23	Sequence 23, Appli
94	24	77.4	215	1	US-08-570-923-23	Sequence 23, Appli
95	24	77.4	215	1	US-08-580-014-23	Sequence 23, Appli
96	24	77.4	215	4	US-09-079-785-23	Sequence 23, Appli
97	24	77.4	218	4	US-08-914-375C-6	Sequence 6, Appli
98	24	77.4	218	4	US-08-914-375C-11	Sequence 11, Appli
99	24	77.4	218	4	US-08-914-375C-12	Sequence 12, Appli
100	24	77.4	218	4	US-08-914-375C-14	Sequence 14, Appli

101	24	77.4	218	4	US-08-914-375C-15	Sequence 15, Appl	174	24	77.4	367	3	US-08-809-2856-2	Sequence 2, Appl
102	24	77.4	218	4	US-08-914-375C-16	Sequence 16, Appl	175	24	77.4	378	2	US-08-707-237N-104	Sequence 104, Appl
103	24	77.4	218	4	US-08-914-375C-30	Sequence 30, Appl	176	24	77.4	378	3	US-08-642-246-26	Sequence 26, Appl
104	24	77.4	218	4	US-09-149-476-717	Sequence 717, App	177	24	77.4	378	4	US-09-451-206-26	Sequence 26, Appl
105	24	77.4	234	1	US-08-225-989-8	Sequence 8, Appl	178	24	77.4	378	5	PCT-US96-06229-26	Sequence 26, Appl
106	24	77.4	234	1	US-08-570-923-8	Sequence 8, Appl	179	24	77.4	387	4	US-09-097-767A-23	Sequence 23, Appl
107	24	77.4	234	1	US-08-580-014-8	Sequence 8, Appl	180	24	77.4	402	1	US-08-378-761A-17	Sequence 17, Appl
108	24	77.4	234	4	US-09-079-785-8	Sequence 8, Appl	181	24	77.4	402	1	US-08-485-286-17	Sequence 17, Appl
109	24	77.4	236	4	US-08-823-120-6	Sequence 16, Appl	182	24	77.4	402	4	US-09-097-767A-26	Sequence 26, Appl
110	24	77.4	238	4	US-09-227-716-16	Sequence 16, Appl	183	24	77.4	400	1	US-08-427-993A-7	Sequence 7, Appl
111	24	77.4	238	4	US-09-609-161B-16	Sequence 16, Appl	184	24	77.4	430	2	US-08-478-609A-3	Sequence 7, Appl
112	24	77.4	240	4	US-08-823-120-3	Sequence 3, Appl	185	24	77.4	430	2	US-09-256-000-23	Sequence 23, Appl
113	24	77.4	251	1	US-08-397-633A-74	Sequence 74, Appl	186	24	77.4	433	4	US-09-342-647-31	Sequence 31, Appl
114	24	77.4	254	1	US-08-378-761A-11	Sequence 11, Appl	187	24	77.4	438	1	US-08-111-939-17	Sequence 17, Appl
115	24	77.4	254	3	US-08-485-286-11	Sequence 11, App	188	24	77.4	438	1	US-09-333-989-9	Sequence 9, Appl
116	24	77.4	254	3	US-08-486-099-114	Sequence 114, App	189	24	77.4	455	1	US-07-872-678A-47	Sequence 47, Appl
117	24	77.4	254	3	US-08-360-107A-124	Sequence 124, App	190	24	77.4	455	4	US-09-347-878A-44	Sequence 44, Appl
118	24	77.4	254	3	US-08-484-223B-114	Sequence 114, App	191	24	77.4	455	4	US-09-356-000-2	Sequence 2, Appl
119	24	77.4	254	3	US-08-919-597-114	Sequence 114, App	192	24	77.4	455	4	US-08-397-633A-78	Sequence 78, Appl
120	24	77.4	254	3	US-08-475-668A-114	Sequence 114, App	193	24	77.4	466	1	US-08-397-633A-77	Sequence 77, Appl
121	24	77.4	254	3	US-08-485-551A-114	Sequence 114, App	194	24	77.4	489	1	US-08-378-761A-19	Sequence 19, Appl
122	24	77.4	254	3	US-08-471-913A-114	Sequence 114, App	195	24	77.4	489	1	US-08-485-286-19	Sequence 19, Appl
123	24	77.4	254	3	US-08-485-264A-114	Sequence 114, App	196	24	77.4	493	1	US-08-378-761A-23	Sequence 23, Appl
124	24	77.4	254	4	US-08-474-349A-114	Sequence 114, App	197	24	77.4	493	1	US-08-485-286-23	Sequence 23, Appl
125	24	77.4	257	1	US-08-378-761A-13	Sequence 13, Appl	198	24	77.4	502	1	US-08-875-761A-25	Sequence 25, Appl
126	24	77.4	257	1	US-08-485-286-13	Sequence 13, Appl	199	24	77.4	502	1	US-08-875-761A-25	Sequence 25, Appl
127	24	77.4	258	1	US-09-097-767A-6	Sequence 6, Appl	200	24	77.4	509	1	US-08-427-993A-1	Sequence 1, Appl
128	24	77.4	261	1	US-08-378-761A-9	Sequence 9, Appl	201	24	77.4	509	2	US-08-478-609A-1	Sequence 8, Appl
129	24	77.4	261	1	US-08-485-286-9	Sequence 9, Appl	202	24	77.4	515	4	US-09-171-461-8	Sequence 8, Appl
130	24	77.4	262	1	US-08-397-633A-73	Sequence 73, Appl	203	24	77.4	517	4	US-08-984-618A-7	Sequence 7, Appl
131	24	77.4	265	1	US-08-378-761A-21	Sequence 21, Appl	204	24	77.4	532	1	US-08-657-192-9	Sequence 9, Appl
132	24	77.4	265	1	US-08-485-286-21	Sequence 21, Appl	205	24	77.4	532	3	US-08-523-373-7	Sequence 7, Appl
133	24	77.4	265	4	US-09-082-649B-82	Sequence 82, Appl	206	24	77.4	533	4	US-09-256-000-19	Sequence 19, Appl
134	24	77.4	271	2	US-07-977-630-5	Sequence 5, Appl	207	24	77.4	537	1	US-08-657-192-15	Sequence 15, Appl
135	24	77.4	271	2	US-08-850-880-6	Sequence 6, Appl	208	24	77.4	556	4	US-09-385-028-9	Sequence 9, Appl
136	24	77.4	271	2	US-08-944-916-6	Sequence 6, Appl	209	24	77.4	570	4	US-09-661-711A-6	Sequence 6, Appl
137	24	77.4	271	2	US-08-814-877-6	Sequence 6, Appl	210	24	77.4	582	4	US-08-913-159-10	Sequence 10, Appl
138	24	77.4	271	4	US-08-760-615-8	Sequence 8, Appl	211	24	77.4	580	4	US-09-367-633A-36	Sequence 36, Appl
139	24	77.4	271	4	US-09-131-028A-5	Sequence 5, Appl	212	24	77.4	583	2	US-08-987-466-1	Sequence 2, Appl
140	24	77.4	271	4	US-09-131-028A-7	Sequence 7, Appl	213	24	77.4	593	4	US-09-240-359-1	Sequence 1, Appl
141	24	77.4	271	4	US-09-272-432A-6	Sequence 6, Appl	214	24	77.4	603	4	US-09-134-007C-5226	Sequence 5226, Ap
142	24	77.4	276	6	5248606-41	Sequence 6, Appl	215	24	77.4	611	4	US-09-370-807-2	Sequence 2, Appl
143	24	77.4	276	1	US-08-378-761A-7	Sequence 7, Appl	216	24	77.4	613	4	US-09-921-259-2	Sequence 2, Appl
144	24	77.4	276	1	US-08-485-286-7	Sequence 7, Appl	217	24	77.4	613	4	US-09-446-504-5	Sequence 5, Appl
145	24	77.4	276	1	US-09-071-035-288	Sequence 288, App	218	24	77.4	613	4	US-09-712-266-5	Sequence 5, Appl
146	24	77.4	278	1	US-07-941-661-2	Sequence 2, Appl	219	24	77.4	613	4	US-09-091-889A-2	Sequence 2, Appl
147	24	77.4	278	1	US-08-279-966-2	Sequence 2, Appl	220	24	77.4	615	3	US-08-860-091A-4	Sequence 4, Appl
148	24	77.4	279	4	US-09-097-767A-10	Sequence 10, Appl	221	24	77.4	615	4	US-09-245-323A-8	Sequence 8, Appl
149	24	77.4	279	4	US-09-097-767A-14	Sequence 14, Appl	222	24	77.4	682	1	US-08-642-255-126	Sequence 126, App
150	24	77.4	280	4	US-09-720-817-1	Sequence 1, Appl	223	24	77.4	682	1	US-08-397-633A-36	Sequence 36, Appl
151	24	77.4	285	6	5248606-45	Sequence 6, Appl	224	24	77.4	717	4	US-09-307-143-2	Sequence 2, Appl
152	24	77.4	285	1	US-08-378-761A-5	Sequence 5, Appl	225	24	77.4	722	1	US-08-481-626-2	Sequence 6, Appl
153	24	77.4	286	1	US-08-485-286-5	Sequence 5, Appl	226	24	77.4	722	1	US-08-533-669A-18	Sequence 18, Appl
154	24	77.4	287	4	US-09-097-767A-17	Sequence 17, Appl	227	24	77.4	732	2	US-08-989-299-4	Sequence 4, Appl
155	24	77.4	290	4	US-09-144-001C-4972	Sequence 4972, Ap	228	24	77.4	732	4	US-08-989-299-4	Sequence 4, Appl
156	24	77.4	300	4	US-09-619-353-5	Sequence 5, Appl	229	24	77.4	732	4	US-09-183-861-18	Sequence 18, Appl
157	24	77.4	301	1	US-08-378-761A-2	Sequence 2, Appl	230	24	77.4	732	4	US-09-022-765-18	Sequence 18, Appl
158	24	77.4	301	1	US-08-485-286-2	Sequence 2, Appl	231	24	77.4	750	4	US-08-806-029-25	Sequence 25, Appl
159	24	77.4	301	6	5248606-2	Sequence 6, Appl	232	24	77.4	751	2	US-08-707-237A-84	Sequence 84, Appl
160	24	77.4	303	4	US-09-071-035-286	Sequence 286, App	233	24	77.4	762	1	US-08-642-255-114	Sequence 114, App
161	24	77.4	315	1	US-08-378-761A-15	Sequence 15, Appl	234	24	77.4	762	1	US-08-642-255-120	Sequence 120, App
162	24	77.4	315	1	US-08-485-286-15	Sequence 15, Appl	235	24	77.4	762	1	US-08-397-633A-26	Sequence 26, Appl
163	24	77.4	315	4	US-09-720-817-3	Sequence 3, Appl	236	24	77.4	762	1	US-08-397-633A-31	Sequence 31, Appl
164	24	77.4	315	4	US-09-720-817-6	Sequence 6, Appl	237	24	77.4	884	1	US-08-397-633A-68	Sequence 68, Appl
165	24	77.4	320	4	US-09-256-000-17	Sequence 17, Appl	238	24	77.4	884	2	US-08-435-641-15	Sequence 15, Appl
166	24	77.4	325	4	US-08-887-534A-53	Sequence 53, Appl	239	24	77.4	884	2	US-08-707-237A-96	Sequence 96, Appl
167	24	77.4	331	3	US-08-793-426A-2	Sequence 2, Appl	240	24	77.4	884	3	US-08-642-246-15	Sequence 15, Appl
168	24	77.4	331	4	US-09-294-565-2	Sequence 2, Appl	241	24	77.4	884	4	US-09-451-206-15	Sequence 15, Appl
169	24	77.4	336	4	US-08-855-910-6	Sequence 6, Appl	242	24	77.4	884	5	PCT-US96-06229-15	Sequence 15, Appl
170	24	77.4	354	1	US-08-447-702-5	Sequence 5, Appl	243	24	77.4	936	2	US-08-707-237A-108	Sequence 108, App
171	24	77.4	354	1	US-08-465-615-5	Sequence 5, Appl	244	24	77.4	936	3	US-08-642-246-30	Sequence 30, Appl
172	24	77.4	356	4	US-08-887-534A-72	Sequence 72, Appl	245	24	77.4	936	4	US-09-451-206-30	Sequence 30, Appl
173	24	77.4	362	4	US-09-134-001C-5403	Sequence 5403, Ap	246	24	77.4	936	5	PCT-US96-06229-30	Sequence 30, Appl

247	24	77.4	946	3	US-09-074-579-3	Sequence 3, Appl1	320	23	74.2	260	4	US-09-187-789-2	Sequence 2, Appl1
248	24	77.4	946	4	US-09-388-774-3	Sequence 3, Appl1	321	23	74.2	260	4	US-09-139-600-2	Sequence 2, Appl1
249	24	77.4	946	5	US-09-546-153-1	Sequence 1, Appl1	322	23	74.2	341	1	US-07-748-783-4	Sequence 4, Appl1
250	24	77.4	966	3	US-08-642-246-34	Sequence 34, Appl	323	23	74.2	341	1	US-08-166-818-4	Sequence 4, Appl1
251	24	77.4	966	4	US-09-451-206-34	Sequence 34, Appl	324	23	74.2	341	1	US-09-134-001C-4042	Sequence 4042, Ap
252	24	77.4	966	5	PCT-US96-06229-34	Sequence 34, Appl	325	23	74.2	350	4	US-09-134-001C-3383	Sequence 3383, Ap
253	24	77.4	1002	3	US-08-707-237A-103	Sequence 103, Appl	326	23	74.2	354	1	US-07-755-556-2	Sequence 2, Appl1
254	24	77.4	1002	3	US-08-642-246-25	Sequence 25, Appl	327	23	74.2	359	4	US-09-155-920-2	Sequence 2, Appl1
255	24	77.4	1002	3	US-09-451-206-25	Sequence 25, Appl	328	23	74.2	359	4	US-08-858-207A-89	Sequence 389, Ap
256	24	77.4	1002	5	PCT-US96-06229-25	Sequence 25, Appl	329	23	74.2	370	4	US-09-134-001C-4393	Sequence 4393, Ap
257	24	77.4	1043	3	US-08-724-354D-4	Sequence 4, Appl1	330	23	74.2	376	4	US-09-134-001C-4259	Sequence 4259, Ap
258	24	77.4	1043	3	US-09-270-984A-4	Sequence 4, Appl1	331	23	74.2	387	4	US-09-364-230-18	Sequence 18, Appl
259	24	77.4	1118	2	US-08-724-354D-2	Sequence 2, Appl1	332	23	74.2	391	4	US-08-429-742-8	Sequence 8, Appl
260	24	77.4	1118	2	US-09-770-984A-2	Sequence 2, Appl1	333	23	74.2	393	1	US-08-429-742-8	Sequence 8, Appl
261	24	77.4	1176	4	US-08-257-999-2	Sequence 7, Appl1	334	23	74.2	397	4	US-09-134-001C-3030	Sequence 3030, Ap
262	24	77.4	1306	4	US-08-989-299-7	Sequence 14, Appl	335	23	74.2	397	4	US-08-860-150-3	Sequence 5504, Ap
263	24	77.4	1872	1	US-08-188-582-14	Sequence 14, Appl	336	23	74.2	404	3	US-09-338-132-3	Sequence 3, Appl1
264	24	77.4	1872	1	US-08-646-715-14	Sequence 14, Appl	337	23	74.2	404	3	US-08-844-064-7	Sequence 7, Appl1
265	24	77.4	1893	1	US-08-188-582-11	Sequence 11, Appl	338	23	74.2	423	3	US-09-009-433-7	Sequence 6, Appl1
266	24	77.4	1893	1	US-08-646-715-11	Sequence 11, Appl	339	23	74.2	423	3	US-09-092-315-6	Sequence 12, Appl
267	24	77.4	3248	1	US-08-353-700-11	Sequence 1, Appl1	340	23	74.2	425	4	US-09-092-315-6	Sequence 12, Appl
268	24	77.4	3248	5	PCT-US95-16216-1	Sequence 1, Appl1	341	23	74.2	435	1	US-08-484-105-12	Sequence 6, Appl1
269	23	74.2	14	2	US-08-372-197-8	Sequence 8, Appl1	342	23	74.2	435	1	US-08-484-106-12	Sequence 12, Appl
270	23	74.2	15	4	US-09-347-504-59	Sequence 59, Appl	343	23	74.2	443	1	US-08-570-157-6	Sequence 6, Appl1
271	23	74.2	16	1	US-08-453-472-10	Sequence 10, Appl	344	23	74.2	443	4	US-09-076-510-6	Sequence 6, Appl1
272	23	74.2	16	1	US-08-038-948-12	Sequence 12, Appl	345	23	74.2	444	1	US-07-937-609-14	Sequence 14, Appl
273	23	74.2	16	1	US-08-453-952-10	Sequence 10, Appl	346	23	74.2	444	4	US-08-029-170-14	Sequence 14, Appl
274	23	74.2	16	1	US-08-862-903-10	Sequence 10, Appl	347	23	74.2	454	4	US-09-233-989-4	Sequence 4, Appl1
275	23	74.2	36	4	US-09-092-315-14	Sequence 14, Appl	348	23	74.2	454	4	US-09-092-315-8	Sequence 8, Appl1
276	23	74.2	40	4	US-09-277-715-25	Sequence 25, Appl	349	23	74.2	464	4	US-09-092-315-1	Sequence 1, Appl1
277	23	74.2	40	4	US-09-609-161B-25	Sequence 25, Appl	350	23	74.2	466	2	US-08-836-791-9	Sequence 9, Appl1
278	23	74.2	54	4	US-08-936-165A-320	Sequence 320, App	351	23	74.2	476	4	US-09-092-315-5	Sequence 5, Appl1
279	23	74.2	56	4	US-09-187-789-64	Sequence 64, Appl	352	23	74.2	478	4	US-09-092-315-7	Sequence 7, Appl1
280	23	74.2	56	4	US-09-139-600-59	Sequence 59, Appl	353	23	74.2	479	4	US-08-985-343-4	Sequence 4, Appl1
281	23	74.2	87	4	US-08-936-165A-406	Sequence 406, App	354	23	74.2	480	2	US-08-272-255-12	Sequence 12, Appl
282	23	74.2	114	4	US-08-858-207A-355	Sequence 355, App	355	23	74.2	480	5	PCT-US95-08665-12	Sequence 12, Appl
283	23	74.2	115	4	US-09-256-000-13	Sequence 13, Appl	356	23	74.2	486	4	US-09-036-987A-16	Sequence 16, Appl
284	23	74.2	137	4	US-09-134-001C-5560	Sequence 5560, Ap	357	23	74.2	486	4	US-09-370-700-16	Sequence 16, Appl
285	23	74.2	138	4	US-09-354-922-1	Sequence 1, Appl1	358	23	74.2	486	4	US-09-091-315-2	Sequence 2, Appl1
286	23	74.2	152	1	US-08-046-583-11	Sequence 11, Appl	359	23	74.2	486	4	US-08-481-435-12	Sequence 12, Appl
287	23	74.2	152	1	US-08-261-677-11	Sequence 11, Appl	360	23	74.2	532	3	US-08-481-435-12	Sequence 12, Appl
288	23	74.2	152	1	US-08-384-556A-2	Sequence 2, Appl1	361	23	74.2	553	3	US-08-481-435-12	Sequence 12, Appl
289	23	74.2	152	1	US-08-384-556A-7	Sequence 7, Appl1	362	23	74.2	571	4	US-09-071-035-16	Sequence 16, Appl
290	23	74.2	152	1	US-08-592-936B-17	Sequence 17, Appl	363	23	74.2	587	1	US-07-935-905A-23	Sequence 23, Appl
291	23	74.2	152	2	US-08-331-355A-11	Sequence 11, Appl	364	23	74.2	593	4	US-09-071-035-16	Sequence 16, Appl
292	23	74.2	152	2	US-08-331-355A-23	Sequence 23, Appl	365	23	74.2	620	4	US-08-461-004A-65	Sequence 65, Appl
293	23	74.2	152	2	US-09-111-573-17	Sequence 17, Appl	366	23	74.2	639	2	US-08-557-309B-37	Sequence 37, Appl
294	23	74.2	152	5	PCT-US94-12364-11	Sequence 11, Appl	367	23	74.2	639	3	US-08-934-306-37	Sequence 37, Appl
295	23	74.2	152	5	PCT-US94-12364-23	Sequence 23, Appl	368	23	74.2	639	4	US-08-934-306-37	Sequence 37, Appl
296	23	74.2	152	5	PCT-US95-07753-2	Sequence 2, Appl1	369	23	74.2	639	4	US-09-256-976-37	Sequence 37, Appl
297	23	74.2	152	5	PCT-US95-07753-7	Sequence 7, Appl1	370	23	74.2	650	1	US-08-325-071-57	Sequence 57, Appl
298	23	74.2	160	4	US-09-134-001C-5659	Sequence 5659, Ap	371	23	74.2	650	4	US-08-461-004A-67	Sequence 67, Appl
299	23	74.2	164	2	US-08-809-267-6	Sequence 6, Appl1	372	23	74.2	680	4	US-08-227-496C-15	Sequence 15, Appl
300	23	74.2	164	5	PCT-US95-1362A-6	Sequence 6, Appl1	373	23	74.2	681	2	US-08-272-255-9	Sequence 25, Appl
301	23	74.2	180	4	US-09-134-001C-4215	Sequence 4215, Ap	374	23	74.2	681	2	US-08-272-255-9	Sequence 25, Appl
302	23	74.2	183	1	US-08-602-010A-4	Sequence 4, Appl1	375	23	74.2	681	3	US-08-964-268-5	Sequence 5, Appl1
303	23	74.2	183	1	US-08-680-726A-4	Sequence 4, Appl1	376	23	74.2	681	5	PCT-US95-08665-6	Sequence 6, Appl1
304	23	74.2	183	4	US-09-092-409-4	Sequence 4, Appl1	377	23	74.2	681	5	PCT-US95-08665-9	Sequence 9, Appl1
305	23	74.2	188	1	US-08-486-715-5	Sequence 5, Appl1	378	23	74.2	684	5	US-09-307-143-4	Sequence 4, Appl1
306	23	74.2	188	1	US-08-486-719-5	Sequence 5, Appl1	379	23	74.2	713	1	US-08-453-472-6	Sequence 6, Appl1
307	23	74.2	188	1	US-08-476-100-5	Sequence 5, Appl1	380	23	74.2	713	1	US-08-038-948-7	Sequence 7, Appl1
308	23	74.2	188	1	US-08-286-767-3	Sequence 3, Appl1	381	23	74.2	713	1	US-08-038-948-8	Sequence 8, Appl1
309	23	74.2	188	1	US-08-475-749-5	Sequence 5, Appl1	382	23	74.2	713	1	US-08-038-948-10	Sequence 10, Appl
310	23	74.2	196	4	US-09-752-165-95	Sequence 95, Appl	383	23	74.2	713	1	US-08-453-472-6	Sequence 6, Appl1
311	23	74.2	218	4	US-08-914-375C-13	Sequence 13, Appl	384	23	74.2	713	2	US-08-862-903-6	Sequence 6, Appl1
312	23	74.2	226	4	US-09-504-358-4	Sequence 4, Appl1	385	23	74.2	732	4	US-09-307-143-4	Sequence 4, Appl1
313	23	74.2	226	4	US-09-954-314-4	Sequence 14, Appl	386	23	74.2	762	2	US-08-907-166-10	Sequence 10, Appl
314	23	74.2	245	2	US-08-272-255-15	Sequence 15, Appl	387	23	74.2	793	1	US-08-188-228-54	Sequence 54, Appl
315	23	74.2	245	5	PCT-US95-08565-15	Sequence 15, Appl	388	23	74.2	793	1	US-08-332-443-48	Sequence 48, Appl
316	23	74.2	246	2	US-08-720-258-2	Sequence 2, Appl1	389	23	74.2	799	1	US-08-332-443-48	Sequence 48, Appl
317	23	74.2	253	2	US-08-720-258-6	Sequence 6, Appl1	390	23	74.2	799	1	US-08-188-228-42	Sequence 42, Appl
318	23	74.2	255	4	US-09-134-001C-3680	Sequence 3680, Ap	391	23	74.2	799	1	US-08-332-443-42	Sequence 42, Appl
319	23	74.2	258	2	US-08-720-258-4	Sequence 4, Appl1	392	23	74.2	823	3	US-08-481-435-4	Sequence 4, Appl1

393	23	74.2	825	4	US-09-540-824-26	Sequence 26, Appl	466	22	71.0	99	4	US-09-188-930-340	Sequence 340, App
394	23	74.2	836	3	US-08-481-435-9	Sequence 9, Appl	467	22	71.0	102	3	US-08-906-768-125	Sequence 125, App
395	23	74.2	844	3	US-08-481-435-7	Sequence 7, Appl	468	22	71.0	102	3	US-08-906-768-125	Sequence 125, App
396	23	74.2	844	3	US-08-481-435-8	Sequence 8, Appl	469	22	71.0	102	4	US-08-639-075A-125	Sequence 125, App
397	23	74.2	876	3	US-08-966-621-2	Sequence 2, Appl	470	22	71.0	102	4	US-09-012-693-135	Sequence 125, App
398	23	74.2	876	3	US-08-966-621-2	Sequence 2, Appl	471	22	71.0	102	4	US-09-012-693-135	Sequence 125, App
399	23	74.2	878	2	US-08-708-541A-26	Sequence 26, Appl	472	22	71.0	102	4	US-08-906-613-135	Sequence 125, App
400	23	74.2	883	2	US-08-953-992-2	Sequence 2, Appl	473	22	71.0	103	2	US-08-480-473B-52	Sequence 52, Appl
401	23	74.2	903	1	US-08-220-151-8	Sequence 8, Appl	474	22	71.0	103	3	US-08-915-213-52	Sequence 52, Appl
402	23	74.2	903	1	US-08-413-118-8	Sequence 8, Appl	475	22	71.0	103	3	US-09-235-217-52	Sequence 52, Appl
403	23	74.2	903	3	US-08-804-439A-22	Sequence 22, Appl	476	22	71.0	110	5	US-08-422-613-2	Sequence 2, Appl
404	23	74.2	903	3	US-08-473-446-8	Sequence 8, Appl	477	22	71.0	110	5	PCT-US96-03916-4	Sequence 4, Appl
405	23	74.2	903	3	US-08-720-229-22	Sequence 22, Appl	478	22	71.0	111	2	US-08-825-556A-2	Sequence 2, Appl
406	23	74.2	904	4	US-08-632-537-1	Sequence 1, Appl	479	22	71.0	111	2	US-08-485-937-4	Sequence 4, Appl
407	23	74.2	904	4	US-08-632-537-2	Sequence 1, Appl	480	22	71.0	111	5	PCT-US93-06552-4	Sequence 4, Appl
408	23	74.2	904	5	PCT-US96-05316-1	Sequence 1, Appl	481	22	71.0	111	5	PCT-US93-06552-4	Sequence 4, Appl
409	23	74.2	904	5	PCT-US96-05316-2	Sequence 2, Appl	482	22	71.0	111	5	PCT-US93-06552-4	Sequence 4, Appl
410	23	74.2	904	6	5244792-3	Patent No. 5244792	483	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
411	23	74.2	904	6	5244792-3	Patent No. 5244792	484	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
412	23	74.2	1052	2	US-08-852-806-2	Sequence 2, Appl	485	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
413	23	74.2	1052	3	US-09-153-669-2	Sequence 2, Appl	486	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
414	23	74.2	1141	1	US-08-353-300-2	Sequence 2, Appl	487	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
415	23	74.2	1153	1	US-08-314-917-2	Sequence 2, Appl	488	22	71.0	126	4	US-09-568-102-12	Sequence 12, Appl
416	23	74.2	1153	1	US-08-265-046-2	Sequence 2, Appl	489	22	71.0	127	3	US-08-705-771-12	Sequence 12, Appl
417	23	74.2	1153	2	US-08-465-522-2	Sequence 2, Appl	490	22	71.0	133	4	US-09-188-930-157	Sequence 157, App
418	23	74.2	1153	5	PCT-US93-11401-2	Sequence 2, Appl	491	22	71.0	135	4	US-09-097-199-84	Sequence 84, Appl
419	23	74.2	1153	5	PCT-US95-07849-2	Sequence 2, Appl	492	22	71.0	135	4	US-09-097-199-84	Sequence 84, Appl
420	23	74.2	1213	1	US-08-188-582-9-2	Sequence 20, Appl	493	22	71.0	139	4	US-08-414-033A-8	Sequence 8, Appl
421	23	74.2	1213	1	US-08-646-715-20	Sequence 20, Appl	494	22	71.0	141	2	US-08-440-894A-8	Sequence 8, Appl
422	23	74.2	1213	1	US-08-680-326-33	Sequence 33, Appl	495	22	71.0	141	2	US-08-485-937-1	Sequence 1, Appl
423	23	74.2	1332	2	US-08-971-244-2	Sequence 2, Appl	496	22	71.0	141	2	US-08-373-215-1	Sequence 1, Appl
424	23	74.2	1332	2	US-09-286-891-2	Sequence 2, Appl	497	22	71.0	141	5	PCT-US93-06552-1	Sequence 523, App
425	23	74.2	1664	1	US-09-599-651-2	Sequence 2, Appl	498	22	71.0	154	4	US-09-446-880A-2	Sequence 2, Appl
426	23	74.2	1664	4	US-08-642-846-2	Sequence 2, Appl	499	22	71.0	155	3	US-09-335-403-20	Sequence 20, Appl
427	23	74.2	1664	4	US-09-264-604-2	Sequence 2, Appl	500	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
428	23	74.2	1820	3	US-07-898-2895-8	Sequence 8, Appl	501	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
429	23	74.2	2100	3	US-08-808-793-23	Sequence 23, Appl	502	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
430	23	74.2	2100	3	US-08-772-512A-19	Sequence 19, Appl	503	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
431	23	74.2	2104	3	US-08-808-793-4	Sequence 4, Appl	504	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
432	23	74.2	2104	3	US-08-772-512A-4	Sequence 4, Appl	505	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
433	23	74.2	2105	3	US-08-808-793-3	Sequence 3, Appl	506	22	71.0	155	4	US-09-568-102-20	Sequence 20, Appl
434	23	74.2	2105	3	US-08-772-512A-3	Sequence 3, Appl	507	22	71.0	163	4	US-09-134-001C-5462	Sequence 5462, Ap
435	23	74.2	2366	1	US-08-480-604A-10	Sequence 10, Appl	508	22	71.0	167	4	US-09-134-001C-4690	Sequence 4690, Ap
436	23	74.2	2366	2	US-08-405-496A-10	Sequence 10, Appl	509	22	71.0	168	3	US-08-669-408B-4	Sequence 3765, Ap
437	23	74.2	2366	4	US-08-915-136-10	Sequence 10, Appl	510	22	71.0	176	2	US-08-989-286-3	Sequence 4, Appl
438	23	74.2	2366	4	US-08-957-310-10	Sequence 10, Appl	511	22	71.0	179	4	US-08-828-533-1	Sequence 1, Appl
439	23	74.2	2366	4	US-09-180-422B-27	Sequence 27, Appl	512	22	71.0	180	4	US-08-828-533-1	Sequence 1, Appl
440	22	71.0	4536	4	US-08-722-126A-15	Sequence 15, Appl	513	22	71.0	188	4	US-09-068-655-5	Sequence 16, Appl
441	22	71.0	11	4	US-08-652-877-64	Sequence 64, Appl	514	22	71.0	204	2	US-08-760-075A-16	Sequence 16, Appl
442	22	71.0	11	4	US-08-476-515A-64	Sequence 64, Appl	515	22	71.0	204	4	US-09-338-546-16	Sequence 16, Appl
443	22	71.0	11	4	US-09-347-926-5	Sequence 3, Appl	516	22	71.0	204	4	US-09-338-546-16	Sequence 16, Appl
444	22	71.0	11	4	US-09-347-926-7	Sequence 7, Appl	517	22	71.0	204	4	US-09-338-546-16	Sequence 16, Appl
445	22	71.0	16	2	US-08-485-937-8	Sequence 8, Appl	518	22	71.0	213	4	US-09-291-170A-5	Sequence 5, Appl
446	22	71.0	16	2	US-08-373-215-8	Sequence 8, Appl	519	22	71.0	213	4	US-09-291-170A-5	Sequence 5, Appl
447	22	71.0	16	5	PCT-US93-06552-8	Sequence 8, Appl	520	22	71.0	215	2	US-08-385-335A-12	Sequence 12, Appl
448	22	71.0	16	5	PCT-US93-06552-8	Sequence 8, Appl	521	22	71.0	215	2	US-08-385-335A-12	Sequence 12, Appl
449	22	71.0	50	4	US-07-662-325A-3	Sequence 3, Appl	522	22	71.0	215	2	US-08-659-206A-5	Sequence 5, Appl
450	22	71.0	50	4	US-08-602-999A-50	Sequence 50, Appl	523	22	71.0	215	2	US-08-659-206A-5	Sequence 5, Appl
451	22	71.0	50	4	US-08-278-665-50	Sequence 50, Appl	524	22	71.0	217	4	US-09-416-488-3	Sequence 3, Appl
452	22	71.0	53	1	US-08-313-050-18	Sequence 18, Appl	525	22	71.0	219	1	US-08-186-529-4	Sequence 4, Appl
453	22	71.0	61	4	US-09-134-001C-3042	Sequence 3042, Ap	526	22	71.0	219	1	US-08-640-386A-4	Sequence 4, Appl
454	22	71.0	65	2	US-08-867-087B-36	Sequence 36, Appl	527	22	71.0	219	1	US-08-751-767A-2	Sequence 2, Appl
455	22	71.0	65	2	US-08-867-087B-37	Sequence 37, Appl	528	22	71.0	222	1	US-08-126-593A-2	Sequence 2, Appl
456	22	71.0	65	4	US-08-858-207A-535	Sequence 535, App	529	22	71.0	222	1	US-08-454-039A-2	Sequence 2, Appl
457	22	71.0	65	4	US-08-858-207A-535	Sequence 535, App	530	22	71.0	228	2	US-08-417-495-27	Sequence 27, Appl
458	22	71.0	68	4	US-09-134-001C-4611	Sequence 4611, Ap	531	22	71.0	228	2	US-08-384-391B-27	Sequence 27, Appl
459	22	71.0	77	4	US-09-188-930-345	Sequence 345, App	532	22	71.0	228	5	PCT-US92-01780-27	Sequence 27, Appl
460	22	71.0	77	4	US-09-188-930-346	Sequence 346, App	533	22	71.0	228	5	PCT-US92-01780-27	Sequence 27, Appl
461	22	71.0	77	4	US-09-724-864-70	Sequence 70, Appl	534	22	71.0	228	5	PCT-US92-01780-27	Sequence 27, Appl
462	22	71.0	77	4	US-09-724-864-70	Sequence 70, Appl	535	22	71.0	237	4	US-09-134-001C-3740	Sequence 3740, Ap
463	22	71.0	95	4	US-09-188-930-344	Sequence 344, App	536	22	71.0	238	4	US-09-134-001C-4745	Sequence 4745, Ap
464	22	71.0	95	4	US-09-724-864-68	Sequence 68, Appl	537	22	71.0	244	3	US-09-003-287-6	Sequence 6, Appl
465	22	71.0	99	2	US-08-825-556A-3	Sequence 3, Appl	538	22	71.0	244	3	US-09-003-287-6	Sequence 6, Appl

539	22	71.0	244	3	US-09-003-287-8	Sequence 8, Appl	612	22	71.0	392	1	US-07-994-133-2	Sequence 2, Appl
540	22	71.0	244	4	US-09-518-986-2	Sequence 2, Appl	613	22	71.0	392	6	5196304-2	Patent No. 5196304
541	22	71.0	245	1	US-07-876-284-2	Sequence 2, Appl	614	22	71.0	397	1	US-07-841-646-27	Sequence 27, Appl
542	22	71.0	245	1	US-08-776-151-9	Sequence 9, Appl	615	22	71.0	397	4	US-09-459-133-2	Sequence 4, Appl
543	22	71.0	245	1	US-09-502-653-12	Sequence 12, Appl	616	22	71.0	397	5	PCT-US91-07635-4	Sequence 5369, Ap
544	22	71.0	246	1	US-07-887-0728-4	Sequence 4, Appl	617	22	71.0	398	4	US-09-134-001C-5369	Sequence 2, Appl
545	22	71.0	246	1	US-08-276-151-7	Sequence 7, Appl	618	22	71.0	409	4	US-09-587-066-2	Sequence 2, Appl
546	22	71.0	246	1	US-08-466-444-4	Sequence 4, Appl	619	22	71.0	410	3	US-08-579-667-2	Sequence 2, Appl
547	22	71.0	246	4	US-09-336-093-5	Sequence 5, Appl	620	22	71.0	413	3	US-08-669-406B-10	Sequence 10, Appl
548	22	71.0	253	1	US-08-265-087-4	Sequence 4, Appl	621	22	71.0	416	3	US-08-554-385-21	Sequence 21, Appl
549	22	71.0	253	1	US-08-621-493-4	Sequence 4, Appl	622	22	71.0	421	1	US-09-093-448-4	Sequence 4, Appl
550	22	71.0	253	2	US-08-684-687-4	Sequence 4, Appl	623	22	71.0	450	2	US-08-818-514-6	Sequence 6, Appl
551	22	71.0	253	2	US-08-965-688-4	Sequence 4, Appl	624	22	71.0	450	2	US-09-115-93A-6	Sequence 6, Appl
552	22	71.0	253	2	US-09-260-173-4	Sequence 4, Appl	625	22	71.0	460	4	US-09-647-540A-2	Sequence 2, Appl
553	22	71.0	254	4	US-09-362-831-3	Sequence 3, Appl	626	22	71.0	466	4	US-09-068-195-25	Sequence 25, Appl
554	22	71.0	254	1	US-08-404-732A-5	Sequence 5, Appl	627	22	71.0	471	1	US-08-257-341-9	Sequence 9, Appl
555	22	71.0	254	1	US-08-404-732A-5	Sequence 7, Appl	628	22	71.0	474	4	US-09-134-001C-3176	Sequence 3176, Ap
556	22	71.0	254	1	US-08-404-732A-7	Sequence 9, Appl	629	22	71.0	479	4	US-09-382-155-27	Sequence 27, Appl
557	22	71.0	254	1	US-08-352-990-2	Sequence 2, Appl	630	22	71.0	490	1	US-08-201-118-7	Sequence 7, Appl
558	22	71.0	254	4	US-09-044-796A-14	Sequence 14, Appl	631	22	71.0	490	2	US-08-238-821B-7	Sequence 7, Appl
559	22	71.0	254	4	US-09-725-460A-14	Sequence 14, Appl	632	22	71.0	490	5	PCT-US95-05744-7	Sequence 7, Appl
560	22	71.0	255	4	US-09-134-001C-4801	Sequence 4801, Ap	633	22	71.0	501	1	US-08-451-715A-10	Sequence 10, Appl
561	22	71.0	256	4	US-08-751-767A-36	Sequence 36, Appl	634	22	71.0	508	4	US-09-134-001C-4570	Sequence 4570, Ap
562	22	71.0	256	2	US-08-751-767A-36	Sequence 3942, Ap	635	22	71.0	511	4	US-08-676-444-42	Sequence 42, Appl
563	22	71.0	284	4	US-08-061-889-2	Sequence 2, Appl	636	22	71.0	518	4	US-09-134-001C-4069	Sequence 4069, Ap
564	22	71.0	284	1	US-08-623-428D-2	Sequence 2, Appl	637	22	71.0	519	2	US-08-751-767A-12	Sequence 12, Appl
565	22	71.0	284	1	US-08-623-428D-2	Sequence 2, Appl	638	22	71.0	522	4	US-08-961-083-110	Sequence 120, Ap
566	22	71.0	284	5	PCT-US94-05378-2	Sequence 2, Appl	639	22	71.0	523	2	US-08-997-080-114	Sequence 114, Ap
567	22	71.0	284	5	PCT-US94-09752-4	Sequence 4, Appl	640	22	71.0	523	2	US-08-997-362-114	Sequence 114, Ap
568	22	71.0	286	3	US-09-120-365-78	Sequence 78, Appl	641	22	71.0	523	4	US-09-095-855-114	Sequence 114, Ap
569	22	71.0	286	4	US-09-515-039-78	Sequence 78, Appl	642	22	71.0	523	4	US-09-324-542-114	Sequence 114, Ap
570	22	71.0	291	4	US-09-134-001C-4895	Sequence 4895, Ap	643	22	71.0	523	4	US-09-205-446-114	Sequence 114, Ap
571	22	71.0	292	1	US-07-952-817-25	Sequence 25, Appl	644	22	71.0	528	6	5475095-2	Patent No. 5475095
572	22	71.0	292	1	US-08-325-562-2	Sequence 2, Appl	645	22	71.0	529	6	5217865-2	Patent No. 5217865
573	22	71.0	294	1	US-08-461-607-13	Sequence 2, Appl	646	22	71.0	530	4	US-08-307-499-29	Sequence 29, Appl
574	22	71.0	295	3	US-08-461-607-13	Sequence 13, Appl	647	22	71.0	530	4	US-09-299-268-29	Sequence 29, Appl
575	22	71.0	295	3	US-09-363-600-13	Sequence 13, Appl	648	22	71.0	530	4	US-09-390-234-18	Sequence 18, Appl
576	22	71.0	301	3	US-08-303-861-21	Sequence 21, Appl	649	22	71.0	535	4	US-09-134-001C-3338	Sequence 3338, Ap
577	22	71.0	301	4	US-09-011-073A-1	Sequence 1, Appl	650	22	71.0	540	2	US-08-368-834-20	Sequence 20, Appl
578	22	71.0	301	4	US-09-230-421-2	Sequence 2, Appl	651	22	71.0	540	2	US-08-751-767A-8	Sequence 8, Appl
579	22	71.0	301	4	US-09-347-504-12	Sequence 12, Appl	652	22	71.0	540	2	US-08-461-722-3	Sequence 3, Appl
580	22	71.0	309	2	US-08-997-080-118	Sequence 118, App	653	22	71.0	540	4	US-08-336-221-3	Sequence 3, Appl
581	22	71.0	309	2	US-08-997-362-118	Sequence 118, App	654	22	71.0	540	4	US-08-336-221-3	Sequence 3, Appl
582	22	71.0	309	4	US-09-095-855-118	Sequence 118, App	655	22	71.0	540	4	US-08-336-221-3	Sequence 3, Appl
583	22	71.0	309	4	US-09-324-542-118	Sequence 118, App	656	22	71.0	540	4	US-08-336-221-3	Sequence 3, Appl
584	22	71.0	309	4	US-09-205-426-118	Sequence 118, App	657	22	71.0	540	5	PCT-US94-06362-4	Sequence 4, Appl
585	22	71.0	316	4	US-09-078-691-2	Sequence 2, Appl	658	22	71.0	540	5	PCT-US94-06362-4	Sequence 4, Appl
586	22	71.0	317	3	US-08-461-607-2	Sequence 2, Appl	659	22	71.0	541	2	US-08-467-822-34	Sequence 34, Appl
587	22	71.0	317	3	US-09-363-600-2	Sequence 2, Appl	660	22	71.0	541	2	US-08-447-154-19	Sequence 19, Appl
588	22	71.0	324	3	US-08-461-607-7	Sequence 7, Appl	661	22	71.0	541	2	US-08-997-080-160	Sequence 160, App
589	22	71.0	324	4	US-09-363-600-7	Sequence 7, Appl	662	22	71.0	541	2	US-08-997-362-160	Sequence 160, App
590	22	71.0	327	2	US-08-997-080-162	Sequence 162, App	663	22	71.0	541	4	US-09-095-855-160	Sequence 160, App
591	22	71.0	327	2	US-08-997-362-162	Sequence 162, App	664	22	71.0	541	4	US-08-432-667-34	Sequence 34, Appl
592	22	71.0	327	4	US-09-095-855-162	Sequence 162, App	665	22	71.0	541	4	US-08-466-248-34	Sequence 34, Appl
593	22	71.0	327	4	US-09-324-542-162	Sequence 162, App	666	22	71.0	541	4	US-09-324-542-160	Sequence 160, App
594	22	71.0	327	4	US-09-095-855-162	Sequence 162, App	667	22	71.0	541	4	US-09-205-446-160	Sequence 160, App
595	22	71.0	330	2	US-08-815-176-1	Sequence 1, Appl	668	22	71.0	541	4	US-09-027-169-2	Sequence 2, Appl
596	22	71.0	330	4	US-09-197-344-1	Sequence 1, Appl	669	22	71.0	541	4	PCT-US93-03027-2	Sequence 2, Appl
597	22	71.0	336	3	US-08-749-816-2	Sequence 2, Appl	670	22	71.0	558	5	PCT-US93-03027-2	Sequence 2, Appl
598	22	71.0	336	3	US-09-144-914-2	Sequence 2, Appl	671	22	71.0	560	1	US-08-052-404-23	Sequence 23, Appl
599	22	71.0	337	4	US-09-134-001C-5550	Sequence 5550, Ap	672	22	71.0	560	1	US-08-479-156-33	Sequence 33, Appl
600	22	71.0	342	2	US-08-483-151-2	Sequence 2, Appl	673	22	71.0	561	1	US-08-052-404-24	Sequence 24, Appl
601	22	71.0	344	4	US-09-108-020-34	Sequence 34, Appl	674	22	71.0	561	1	US-08-479-156-34	Sequence 34, Appl
602	22	71.0	347	4	US-09-134-001C-3298	Sequence 3298, Ap	675	22	71.0	564	4	US-09-211-704A-8	Sequence 8, Appl
603	22	71.0	349	3	US-08-806-597A-14	Sequence 14, Appl	676	22	71.0	574	4	US-09-142-623-13	Sequence 13, Appl
604	22	71.0	349	3	US-08-970-428A-14	Sequence 14, Appl	677	22	71.0	579	3	US-08-704-711A-1	Sequence 1, Appl
605	22	71.0	355	4	US-09-134-001C-4752	Sequence 4752, Ap	678	22	71.0	579	4	US-09-521-220-1	Sequence 1, Appl
606	22	71.0	356	1	US-07-959-941-2	Sequence 2, Appl	679	22	71.0	582	4	US-08-704-711A-2	Sequence 2, Appl
607	22	71.0	356	1	US-08-259-924-2	Sequence 2, Appl	680	22	71.0	582	4	US-08-448-489-1	Sequence 1, Appl
608	22	71.0	359	4	US-09-134-001C-4842	Sequence 4842, Ap	681	22	71.0	582	4	US-09-211-704A-9	Sequence 9, Appl
609	22	71.0	364	4	US-09-338-671-2	Sequence 2, Appl	682	22	71.0	582	4	US-09-521-220-2	Sequence 2, Appl
610	22	71.0	367	3	US-09-141-047-10	Sequence 10, Appl	683	22	71.0	584	1	US-09-391-104-18	Sequence 18, Appl
611	22	71.0	375	4	US-09-134-001C-5079	Sequence 5079, Ap	684	22	71.0	593	6	US-08-313-288B-17	Sequence 17, Appl

685	22	71.0	593	6	5523211-2	Patent No. 5523211	758	22	71.0	868	1	US-08-251-937A-6	Sequence 6, Appl
686	22	71.0	595	6	5523211-3	Patent No. 5523211	759	22	71.0	869	1	US-08-212-133A-3	Sequence 3, Appl
687	22	71.0	596	4	US-09-000-145-2	Sequence 2, Appl	760	22	71.0	870	4	US-09-177-650-91	Sequence 91, Appl
688	22	71.0	608	4	US-09-095-385-4	Sequence 4, Appl	761	22	71.0	872	4	US-09-177-650-7	Sequence 7, Appl
689	22	71.0	609	1	US-07-798-776-2	Sequence 2, Appl	762	22	71.0	874	3	US-08-823-110-1	Sequence 1, Appl
690	22	71.0	609	3	US-08-251-288A-2	Sequence 2, Appl	763	22	71.0	908	3	US-08-604-299-1	Sequence 1, Appl
691	22	71.0	609	3	US-09-298-819A-2	Sequence 2, Appl	764	22	71.0	908	3	US-08-699-1103B-12	Sequence 12, Appl
692	22	71.0	609	3	US-09-586-563C-2	Sequence 2, Appl	765	22	71.0	908	4	US-08-855-146-2	Sequence 2, Appl
693	22	71.0	609	4	US-09-586-562C-2	Sequence 2, Appl	766	22	71.0	908	4	US-09-229-051-12	Sequence 12, Appl
694	22	71.0	613	4	US-08-622-740-6	Sequence 6, Appl	767	22	71.0	912	4	US-08-617-785-2	Sequence 2, Appl
695	22	71.0	613	3	US-08-440-689-6	Sequence 6, Appl	768	22	71.0	912	4	US-09-641-318-2	Sequence 2, Appl
696	22	71.0	613	4	US-09-122-399-6	Sequence 11, Appl	769	22	71.0	912	5	PCT-US81-09442-19	Sequence 19, Appl
697	22	71.0	613	4	US-08-447-985-11	Sequence 11, Appl	770	22	71.0	918	2	US-08-843-530B-35	Sequence 35, Appl
698	22	71.0	614	3	US-08-622-740-8	Sequence 8, Appl	771	22	71.0	933	3	US-08-313-200-1	Sequence 1, Appl
699	22	71.0	614	3	US-08-440-689-8	Sequence 8, Appl	772	22	71.0	933	5	PCT-US93-03837-1	Sequence 1, Appl
700	22	71.0	614	4	US-09-122-399-8	Sequence 8, Appl	773	22	71.0	934	1	US-08-446-486-7	Sequence 7, Appl
701	22	71.0	615	4	US-08-447-985-13	Sequence 13, Appl	774	22	71.0	934	1	US-08-463-308-7	Sequence 7, Appl
702	22	71.0	615	3	US-09-042-426-10	Sequence 10, Appl	775	22	71.0	969	1	US-07-671-817A-4	Sequence 4, Appl
703	22	71.0	615	4	US-09-291-238-10	Sequence 10, Appl	776	22	71.0	969	1	US-07-671-817A-6	Sequence 6, Appl
704	22	71.0	615	4	US-09-330-760-10	Sequence 10, Appl	777	22	71.0	979	4	US-08-878-474-5	Sequence 5, Appl
705	22	71.0	615	4	US-09-338-473-10	Sequence 10, Appl	778	22	71.0	990	4	US-09-627-376-7	Sequence 7, Appl
706	22	71.0	615	4	US-09-330-737-10	Sequence 10, Appl	779	22	71.0	993	4	US-08-836-687B-30	Sequence 30, Appl
707	22	71.0	615	4	US-09-329-169-10	Sequence 10, Appl	780	22	71.0	1017	4	US-09-134-001C-3542	Sequence 3542, Ap
708	22	71.0	615	4	US-09-330-714A-10	Sequence 10, Appl	781	22	71.0	1040	4	US-08-661-083C-118	Sequence 118, Ap
709	22	71.0	615	4	US-09-328-826-10	Sequence 10, Appl	782	22	71.0	1042	4	US-09-387-695-2	Sequence 2, Appl
710	22	71.0	617	5	US-08-729-601A-2	Sequence 2, Appl	783	22	71.0	1065	3	US-08-630-172-9	Sequence 9, Appl
711	22	71.0	623	5	PCT-US96-03916-64	Sequence 64, Appl	784	22	71.0	1065	4	US-09-630-172-9	Sequence 9, Appl
712	22	71.0	625	4	US-08-959-004-10	Sequence 10, Appl	785	22	71.0	1090	6	US-09-375-419-9	Sequence 6, Appl
713	22	71.0	648	3	US-08-810-720-13	Sequence 13, Appl	786	22	71.0	1129	6	PCT-US93-03275-6	Sequence 6, Appl
714	22	71.0	649	4	US-08-800-291B-5	Sequence 5, Appl	787	22	71.0	1155	1	US-08-349-867-19	Sequence 19, Appl
715	22	71.0	649	4	US-08-800-291B-6	Sequence 6, Appl	788	22	71.0	1155	1	US-08-349-867-33	Sequence 33, Appl
716	22	71.0	649	4	US-09-066-047-5	Sequence 5, Appl	789	22	71.0	1155	1	US-08-339-476-19	Sequence 19, Appl
717	22	71.0	650	4	US-08-800-291B-4	Sequence 4, Appl	790	22	71.0	1155	1	US-08-446-486-2	Sequence 2, Appl
718	22	71.0	657	4	US-09-370-368-7	Sequence 7, Appl	791	22	71.0	1155	1	US-07-951-715A-9	Sequence 9, Appl
719	22	71.0	669	3	US-08-704-711A-3	Sequence 3, Appl	792	22	71.0	1155	2	US-08-463-308-2	Sequence 2, Appl
720	22	71.0	669	4	US-09-521-220-3	Sequence 3, Appl	793	22	71.0	1155	2	US-08-598-305A-19	Sequence 19, Appl
721	22	71.0	669	4	US-09-391-104-19	Sequence 29, Appl	794	22	71.0	1155	2	US-08-598-305A-33	Sequence 33, Appl
722	22	71.0	677	4	US-08-522-269B-3	Sequence 3, Appl	795	22	71.0	1155	2	US-08-639-923A-19	Sequence 19, Appl
723	22	71.0	677	4	US-09-294-923-3	Sequence 3, Appl	796	22	71.0	1155	2	US-08-459-448A-9	Sequence 9, Appl
724	22	71.0	688	1	US-07-141-047-8	Sequence 8, Appl	797	22	71.0	1155	3	US-08-459-448A-9	Sequence 9, Appl
725	22	71.0	695	1	US-07-671-817A-5	Sequence 5, Appl	798	22	71.0	1155	3	US-09-021-203-2	Sequence 2, Appl
726	22	71.0	720	4	US-09-480-921B-8	Sequence 8, Appl	799	22	71.0	1155	3	US-08-459-504B-9	Sequence 9, Appl
727	22	71.0	724	3	US-09-121-964-1	Sequence 1, Appl	800	22	71.0	1155	3	US-08-459-504A-9	Sequence 9, Appl
728	22	71.0	742	4	US-09-215-664-12	Sequence 12, Appl	801	22	71.0	1155	3	US-09-053-549-4	Sequence 4, Appl
729	22	71.0	743	2	US-09-012-030-2	Sequence 2, Appl	802	22	71.0	1155	3	US-09-053-549-6	Sequence 6, Appl
730	22	71.0	743	2	US-08-350-454-2	Sequence 2, Appl	803	22	71.0	1155	4	US-09-547-422-9	Sequence 9, Appl
731	22	71.0	746	3	US-08-434-000A-4	Sequence 4, Appl	804	22	71.0	1155	5	PCT-US95-05431-19	Sequence 19, Appl
732	22	71.0	746	4	US-09-312-157-4	Sequence 4, Appl	805	22	71.0	1155	6	5254799-4	Patent No. 5254799
733	22	71.0	771	3	US-09-121-964-9	Sequence 9, Appl	806	22	71.0	1156	4	US-09-178-052-15	Sequence 15, Appl
734	22	71.0	815	1	US-08-123-520C-9	Sequence 9, Appl	807	22	71.0	1156	1	US-07-828-788A-14	Sequence 14, Appl
735	22	71.0	817	1	US-08-381-931B-2	Sequence 2, Appl	808	22	71.0	1156	1	US-08-239-476-33	Sequence 33, Appl
736	22	71.0	820	1	US-08-291-896-2	Sequence 2, Appl	809	22	71.0	1156	1	US-08-356-034-4	Sequence 4, Appl
737	22	71.0	820	2	US-08-485-278-2	Sequence 2, Appl	810	22	71.0	1156	2	US-08-639-923A-33	Sequence 33, Appl
738	22	71.0	827	4	US-08-659-286-11	Sequence 11, Appl	811	22	71.0	1156	3	US-08-933-891-4	Sequence 4, Appl
739	22	71.0	827	4	US-09-469-253-11	Sequence 11, Appl	812	22	71.0	1156	4	US-09-176-320-6	Sequence 6, Appl
740	22	71.0	827	4	US-09-642-146-11	Sequence 11, Appl	813	22	71.0	1156	5	PCT-US92-11337-14	Sequence 14, Appl
741	22	71.0	838	4	US-09-315-794-62	Sequence 52, Appl	814	22	71.0	1156	5	PCT-US95-05431-33	Sequence 33, Appl
742	22	71.0	838	4	US-09-389-341-52	Sequence 52, Appl	815	22	71.0	1156	6	5188960-4	Patent No. 5188960
743	22	71.0	838	4	US-09-564-805-229	Sequence 229, App	816	22	71.0	1156	2	US-08-789-078-2	Sequence 2, Appl
744	22	71.0	844	1	US-07-646-537B-2	Sequence 2, Appl	817	22	71.0	1156	2	US-08-752-633-2	Sequence 2, Appl
745	22	71.0	853	3	US-08-451-946B-6	Sequence 6, Appl	818	22	71.0	1156	2	US-08-476-062A-42	Sequence 42, Appl
746	22	71.0	853	3	US-08-446-938B-6	Sequence 6, Appl	819	22	71.0	1156	5	PCT-US95-04886-2	Sequence 2, Appl
747	22	71.0	853	3	US-08-311-703A-6	Sequence 6, Appl	820	22	71.0	1156	5	PCT-US96-01314-42	Sequence 42, Appl
748	22	71.0	853	3	US-08-446-939B-6	Sequence 6, Appl	821	22	71.0	1156	1	US-07-828-788A-6	Sequence 6, Appl
749	22	71.0	853	3	US-09-183-543-6	Sequence 6, Appl	822	22	71.0	1156	1	US-08-278-685-2	Sequence 2, Appl
750	22	71.0	853	3	US-08-699-103B-10	Sequence 10, Appl	823	22	71.0	1156	1	US-08-277-721-2	Sequence 2, Appl
751	22	71.0	853	4	US-08-446-936A-6	Sequence 6, Appl	824	22	71.0	1156	1	US-08-277-721-4	Sequence 4, Appl
752	22	71.0	853	4	US-09-229-059-10	Sequence 10, Appl	825	22	71.0	1156	1	US-08-602-737-4	Sequence 4, Appl
753	22	71.0	853	5	PCT-US92-09326-2	Sequence 2, Appl	826	22	71.0	1156	1	US-09-001-862-4	Sequence 4, Appl
754	22	71.0	854	4	US-09-105-058C-27	Sequence 27, Appl	827	22	71.0	1156	1	PCT-US92-11337-6	Sequence 6, Appl
755	22	71.0	863	4	US-09-238-303-11	Sequence 11, Appl	828	22	71.0	1156	1	US-07-828-788A-12	Sequence 12, Appl
756	22	71.0	866	4	US-09-134-001C-4930	Sequence 4930, Ap	829	22	71.0	1156	1	US-08-446-486-6	Sequence 6, Appl
757	22	71.0	868	1	US-07-864-004B-6	Sequence 6, Appl	830	22	71.0	1156	1	US-08-157-563A-10	Sequence 10, Appl

831	22	71.0	1176	1	US-08-356-034-2	Sequence 2, Appl	904	22	71.0	1181	3	US-08-459-444-13	Sequence 13, Appl
832	22	71.0	1176	1	US-08-463-308-6	Sequence 6, Appl	905	22	71.0	1181	3	US-08-459-444-15	Sequence 15, Appl
833	22	71.0	1176	1	US-08-933-891-2	Sequence 2, Appl	906	22	71.0	1181	3	US-08-459-444-17	Sequence 17, Appl
834	22	71.0	1176	5	PCT-US92-11337-12	Sequence 12, Appl	907	22	71.0	1181	3	US-08-459-444-18	Sequence 18, Appl
835	22	71.0	1176	5	PCT-US93-11405A-10	Sequence 10, Appl	908	22	71.0	1181	4	US-09-547-422-11	Sequence 11, Appl
836	22	71.0	1176	5	PCT-US93-11405A-10	Sequence 10, Appl	909	22	71.0	1181	4	US-09-547-422-13	Sequence 13, Appl
837	22	71.0	1177	1	US-07-828-088A-8	Sequence 8, Appl	910	22	71.0	1181	4	US-09-547-422-15	Sequence 15, Appl
838	22	71.0	1177	1	US-07-920-085-2	Sequence 2, Appl	911	22	71.0	1181	4	US-09-547-422-17	Sequence 17, Appl
839	22	71.0	1177	3	US-08-754-490-10	Sequence 10, Appl	912	22	71.0	1182	4	US-09-547-422-28	Sequence 28, Appl
840	22	71.0	1177	3	US-08-754-490-12	Sequence 12, Appl	913	22	71.0	1182	4	US-08-349-867-34	Sequence 34, Appl
841	22	71.0	1177	3	US-08-754-490-14	Sequence 14, Appl	914	22	71.0	1182	1	US-08-598-305A-34	Sequence 34, Appl
842	22	71.0	1177	3	US-08-754-490-26	Sequence 26, Appl	915	22	71.0	1182	6	5254799-6	Sequence 6, Appl
843	22	71.0	1177	3	US-08-754-490-28	Sequence 28, Appl	916	22	71.0	1186	1	US-08-602-737-6	Sequence 6, Appl
844	22	71.0	1177	3	US-08-855-160-2	Sequence 2, Appl	917	22	71.0	1186	4	US-09-001-982-6	Sequence 6, Appl
845	22	71.0	1177	3	US-08-855-160-6	Sequence 6, Appl	918	22	71.0	1186	4	US-08-239-476-34	Sequence 34, Appl
846	22	71.0	1177	3	US-08-855-160-8	Sequence 8, Appl	919	22	71.0	1188	2	US-08-639-923A-34	Sequence 34, Appl
847	22	71.0	1177	3	US-08-922-505A-10	Sequence 10, Appl	920	22	71.0	1188	5	PCT-US93-03431-34	Sequence 34, Appl
848	22	71.0	1177	3	US-08-922-505A-12	Sequence 12, Appl	921	22	71.0	1188	6	5254799-7	Sequence 6, Appl
849	22	71.0	1177	3	US-08-922-505A-14	Sequence 14, Appl	922	22	71.0	1193	1	US-08-602-737-8	Sequence 8, Appl
850	22	71.0	1177	3	US-08-922-505A-26	Sequence 26, Appl	923	22	71.0	1193	3	US-08-754-490-30	Sequence 30, Appl
851	22	71.0	1177	3	US-08-922-505A-28	Sequence 28, Appl	924	22	71.0	1193	3	US-08-922-505A-30	Sequence 30, Appl
852	22	71.0	1177	3	US-08-922-505A-34	Sequence 34, Appl	925	22	71.0	1193	4	US-09-001-982-8	Sequence 8, Appl
853	22	71.0	1177	4	US-09-260-952A-10	Sequence 10, Appl	926	22	71.0	1193	4	US-09-260-952A-30	Sequence 30, Appl
854	22	71.0	1177	4	US-09-260-952A-12	Sequence 12, Appl	927	22	71.0	1193	4	US-09-263-341-30	Sequence 30, Appl
855	22	71.0	1177	4	US-09-260-952A-14	Sequence 14, Appl	928	22	71.0	1193	4	US-09-263-341-30	Sequence 30, Appl
856	22	71.0	1177	4	US-09-260-952A-26	Sequence 26, Appl	929	22	71.0	1193	4	US-09-263-341A-30	Sequence 30, Appl
857	22	71.0	1177	4	US-09-260-952A-28	Sequence 28, Appl	930	22	71.0	1220	3	US-08-930-996A-2	Sequence 2, Appl
858	22	71.0	1177	4	US-09-253-341-10	Sequence 10, Appl	931	22	71.0	1227	2	US-08-760-075A-18	Sequence 18, Appl
859	22	71.0	1177	4	US-09-253-341-12	Sequence 12, Appl	932	22	71.0	1227	4	US-09-338-546-18	Sequence 18, Appl
860	22	71.0	1177	4	US-09-253-341-14	Sequence 14, Appl	933	22	71.0	1227	4	US-09-338-546-18	Sequence 18, Appl
861	22	71.0	1177	4	US-09-253-341-26	Sequence 26, Appl	934	22	71.0	1227	4	US-09-659-004-18	Sequence 18, Appl
862	22	71.0	1177	4	US-09-253-341-28	Sequence 28, Appl	935	22	71.0	1229	3	US-09-510-293-2	Sequence 2, Appl
863	22	71.0	1177	4	US-09-253-341-34	Sequence 34, Appl	936	22	71.0	1229	4	US-09-510-293-2	Sequence 2, Appl
864	22	71.0	1177	4	US-09-253-331A-10	Sequence 10, Appl	937	22	71.0	1259	4	US-09-040-001C-3757	Sequence 3757, Ap
865	22	71.0	1177	4	US-09-253-331A-12	Sequence 12, Appl	938	22	71.0	1259	4	US-09-134-001C-3757	Sequence 3757, Ap
866	22	71.0	1177	4	US-09-253-331A-14	Sequence 14, Appl	939	22	71.0	1337	5	PCT-US95-05112-2	Sequence 2, Appl
867	22	71.0	1177	4	US-09-253-331A-26	Sequence 26, Appl	940	22	71.0	1337	5	US-08-854-855-2	Sequence 2, Appl
868	22	71.0	1177	4	US-09-253-331A-28	Sequence 28, Appl	941	22	71.0	1337	5	PCT-US95-05112-2	Sequence 2, Appl
869	22	71.0	1177	4	US-09-253-331A-34	Sequence 34, Appl	942	22	71.0	1356	4	US-09-770-170-6	Sequence 2, Appl
870	22	71.0	1177	4	US-09-261-040-12	Sequence 12, Appl	943	22	71.0	1455	2	US-08-726-012B-2	Sequence 2, Appl
871	22	71.0	1177	4	US-09-261-040-14	Sequence 14, Appl	944	22	71.0	1463	4	US-08-157-005-3	Sequence 3, Appl
872	22	71.0	1177	4	US-09-261-040-26	Sequence 26, Appl	945	22	71.0	1463	4	US-08-747-863-3	Sequence 3, Appl
873	22	71.0	1177	4	US-09-261-040-28	Sequence 28, Appl	946	22	71.0	1463	4	US-09-565-864-3	Sequence 3, Appl
874	22	71.0	1177	5	PCT-US92-11337-8	Sequence 8, Appl	947	22	71.0	2071	4	US-09-415-552-6	Sequence 6, Appl
875	22	71.0	1177	6	5169629-2	Sequence 2, Appl	948	22	71.0	2482	1	US-08-328-254-6	Sequence 6, Appl
876	22	71.0	1178	1	US-08-463-308-5	Sequence 5, Appl	949	22	71.0	2860	2	US-08-828-267-2	Sequence 2, Appl
877	22	71.0	1178	1	US-08-463-308-5	Sequence 5, Appl	950	22	71.0	7	7	US-08-438-475-5	Sequence 5, Appl
878	22	71.0	1178	1	US-08-463-308-5	Sequence 5, Appl	951	22	71.0	7	5	PCT-US93-03970-1	Sequence 1, Appl
879	22	71.0	1179	6	5254799-3	Sequence 3, Appl	952	22	71.0	8	1	US-08-438-475-1	Sequence 1, Appl
880	22	71.0	1179	6	5254799-3	Sequence 3, Appl	953	22	71.0	8	5	PCT-US93-03970-1	Sequence 1, Appl
881	22	71.0	1179	6	5254799-3	Sequence 3, Appl	954	22	71.0	8	5	PCT-US93-03970-1	Sequence 1, Appl
882	22	71.0	1179	6	5254799-3	Sequence 3, Appl	955	22	71.0	8	5	PCT-US93-03970-1	Sequence 1, Appl
883	22	71.0	1181	1	US-07-951-715A-11	Sequence 11, Appl	956	22	71.0	9	3	US-08-438-475-2	Sequence 2, Appl
884	22	71.0	1181	1	US-07-951-715A-13	Sequence 13, Appl	957	22	71.0	9	3	US-08-438-475-2	Sequence 2, Appl
885	22	71.0	1181	1	US-07-951-715A-15	Sequence 15, Appl	958	22	71.0	9	4	US-09-452-142-1	Sequence 1, Appl
886	22	71.0	1181	1	US-07-951-715A-17	Sequence 17, Appl	959	22	71.0	9	5	PCT-US93-03970-2	Sequence 2, Appl
887	22	71.0	1181	1	US-07-951-715A-28	Sequence 28, Appl	960	22	71.0	10	4	US-09-130-365-20	Sequence 20, Appl
888	22	71.0	1181	2	US-08-459-448A-11	Sequence 11, Appl	961	22	71.0	10	4	US-09-130-365-20	Sequence 20, Appl
889	22	71.0	1181	2	US-08-459-448A-13	Sequence 13, Appl	962	22	71.0	10	4	US-09-130-365-20	Sequence 20, Appl
890	22	71.0	1181	2	US-08-459-448A-15	Sequence 15, Appl	963	22	71.0	11	4	US-08-652-877-68	Sequence 68, Appl
891	22	71.0	1181	2	US-08-459-448A-17	Sequence 17, Appl	964	22	71.0	11	4	US-08-476-515A-68	Sequence 68, Appl
892	22	71.0	1181	2	US-08-459-448A-28	Sequence 28, Appl	965	22	71.0	11	4	US-09-347-926-2	Sequence 2, Appl
893	22	71.0	1181	2	US-08-459-448A-28	Sequence 28, Appl	966	22	71.0	15	4	US-09-347-926-2	Sequence 2, Appl
894	22	71.0	1181	3	US-08-459-448A-11	Sequence 11, Appl	967	22	71.0	15	4	US-09-347-926-2	Sequence 2, Appl
895	22	71.0	1181	3	US-08-459-448A-13	Sequence 13, Appl	968	22	71.0	16	3	US-08-202-047-1	Sequence 36, Appl
896	22	71.0	1181	3	US-08-459-448A-15	Sequence 15, Appl	969	22	71.0	16	3	US-08-964-990-1	Sequence 1, Appl
897	22	71.0	1181	3	US-08-459-448A-17	Sequence 17, Appl	970	22	71.0	17	4	US-09-120-365-18	Sequence 18, Appl
898	22	71.0	1181	3	US-08-459-448A-28	Sequence 28, Appl	971	22	71.0	17	4	US-08-974-549A-54	Sequence 54, Appl
899	22	71.0	1181	3	US-08-459-448A-11	Sequence 11, Appl	972	22	71.0	18	3	US-09-515-039-18	Sequence 18, Appl
900	22	71.0	1181	3	US-08-459-448A-13	Sequence 13, Appl	973	22	71.0	18	4	US-09-120-365-21	Sequence 21, Appl
901	22	71.0	1181	3	US-08-459-448A-15	Sequence 15, Appl	974	22	71.0	20	1	US-09-515-039-21	Sequence 21, Appl
902	22	71.0	1181	3	US-08-459-448A-17	Sequence 17, Appl	975	22	71.0	20	1	US-08-023-760A-2	Sequence 2, Appl
903	22	71.0	1181	3	US-08-459-448A-28	Sequence 28, Appl	976	22	71.0	27	6	5378464-6	Sequence 6, Appl

```
977 21 67.7 32 2 US-08-140-137A-46 Sequence 46, Appl
978 21 67.7 43 1 US-07-998-003A-73 Sequence 73, Appl
979 21 67.7 43 1 US-08-453-274B-73 Sequence 73, Appl
980 21 67.7 43 1 US-08-453-695A-73 Sequence 73, Appl
981 21 67.7 43 1 US-08-268-161A-73 Sequence 73, Appl
982 21 67.7 43 2 US-08-453-702A-73 Sequence 73, Appl
983 21 67.7 43 4 US-09-099-639-73 Sequence 73, Appl
984 21 67.7 43 5 PCT-US93-12588-73 Sequence 73, Appl
985 21 67.7 43 5 PCT-US93-08071-73 Sequence 73, Appl
986 21 67.7 52 2 US-08-209-521-16 Sequence 16, Appl
987 21 67.7 52 4 US-08-961-810-126 Sequence 126, App
988 21 67.7 52 4 US-08-352-902D-126 Sequence 126, App
989 21 67.7 53 4 US-08-851-843A-20 Sequence 20, Appl
990 21 67.7 53 4 US-08-854-050-20 Sequence 20, Appl
991 21 67.7 53 4 US-09-430-323-20 Sequence 20, Appl
992 21 67.7 59 2 US-08-633-879C-18 Sequence 18, Appl
993 21 67.7 61 6 5320958-19 Patent No. 5320958
994 21 67.7 64 2 US-08-209-521-14 Sequence 14, Appl
995 21 67.7 64 2 US-08-209-521-15 Sequence 15, Appl
996 21 67.7 64 4 US-08-961-810-124 Sequence 124, App
997 21 67.7 64 4 US-08-961-810-125 Sequence 125, App
998 21 67.7 64 4 US-08-352-902D-124 Sequence 124, App
999 21 67.7 64 4 US-08-352-902D-125 Sequence 125, App
1000 21 67.7 69 4 US-09-134-001C-4746 Sequence 4746, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-045-284A-9
; Sequence 9, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107U51
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-9
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 6;
Pred. No. 1.9e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-2
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 274 VRYEDL 279

RESULT 3
US-09-190-911-1
; Sequence 1, Application US/09190911
; Patent No. 6365365
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Tangemann, Kirsten
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CIP
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-190-911-1
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 274 VRYEDL 279

RESULT 4
US-09-015-188-2
; Sequence 2, Application US/09015188C
; Patent No. 6399358
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J
; APPLICANT: Tabas, Ira
; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
; FILE REFERENCE: JEPF-0231
; CURRENT APPLICATION NUMBER: US/09/015,188C
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-015-188-2
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 411;
Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 300 VRYEDL 305
```

RESULT 5
US-09-263-023-2
Sequence 2, Application US/09263023
Patent No. 6037159

GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Habuchi, Reiji
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263,023
CURRENT FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-263-023-2

Query Match
Best Local Similarity 100.0%; Score 31; DB 3; Length 483;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 376 VRYEDL 381

RESULT 6
US-09-471-867-2
Sequence 2, Application US/09471867
Patent No. 6455289

GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Habuchi, Reiji
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471,867
CURRENT FILING DATE: 1999-12-23
EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-471-867-2

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 483;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 376 VRYEDL 381

RESULT 7
US-09-263-023-4
Sequence 4, Application US/09263023
Patent No. 6037159

GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Habuchi, Reiji
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263,023
CURRENT FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-263-023-4

Query Match
Best Local Similarity 100.0%; Score 31; DB 3; Length 484;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 377 VRYEDL 382

RESULT 8
US-09-471-867-4
Sequence 4, Application US/09471867
Patent No. 6455289

GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Habuchi, Reiji
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471,867
CURRENT FILING DATE: 1999-12-23
EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-471-867-4

Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 484;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 377 VRYEDL 382

RESULT 9

US-08-867-030B-13
; Sequence 13, Application US/08867030B
; Patent No. 5948900
; GENERAL INFORMATION:
; APPLICANT: Yotter et al.
; TITLE OF INVENTION: Streptococcus pneumoniae
; TITLE OF INVENTION: Capsular Polysaccharide Genes and Flanking Regions
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch floppy
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,030B
; FILING DATE: June 2, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/243,546
; FILING DATE: May 16, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5923
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; IMMEDIATE SOURCE:
; POSITION IN GENE:
US-08-867-030B-13

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 165 VRYEDV 170

RESULT 10
PCT-US95-06119-13
; Sequence 13, Application PC/TUS9506119
; GENERAL INFORMATION:
; APPLICANT:

TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE CAPSULAR
; TITLE OF INVENTION: POLYSACCHARIDE GENES AND FLANKING REGIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent Release #1.0, Version
SOFTWARE: #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06119
FILING DATE: CONCURRENTLY HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,546
FILING DATE: 16-MAY-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: AMCY018P--

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 306 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-06119-13

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 165 VRYEDV 170

RESULT 11

US-08-655-878-2
; Sequence 2, Application US/08655878
; Patent No. 5827713

GENERAL INFORMATION:

APPLICANT: FUKUTA, MASAKAZU

APPLICANT: HABUCHI, OSAMI

TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE:

STREET:

CITY:

STATE:

COUNTRY:

ZIP:

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/655,878

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 458
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-878-2

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 348 VRYEDV 353

RESULT 12
US-08-899-514-2
Sequence 2, Application US/08899514
Patent No. 5910581
GENERAL INFORMATION:
APPLICANT: HABUCHI, OSAMI
APPLICANT: FUKUTA, MASAKAZU
TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
TITLE OF INVENTION: SULFONASEPASE ORIGINATING FROM HUMAN AND DNA CODING
TITLE OF INVENTION: FOR THE POLYPEPTIDE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBE, MARTENS, OLSON & BEAR, LLP
STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CALIFORNIA
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,514
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: DANIEL E. ALTMAN
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: TOYAM21.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714 760 0404
TELEFAX: 714 760 9502
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 479
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-514-2

Query Match 90.3%; Score 28; DB 2; Length 479;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 369 VRYEDV 374

RESULT 13
US-08-472-934-6
Sequence 6, Application US/08472934
Patent No. 5753446
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lathive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,934
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,421
FILING DATE: 15-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 5,405,941
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr. Esq.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: CPT-004DVC22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-934-6

Query Match 90.3%; Score 28; DB 1; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

Db 64 VRYEDV 69

RESULT 14
US-08-323-460A-6
Sequence 6, Application US/08323460A

Patent No. 5854043

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

STREET: 1700 LINCOLN STREET, SUITE 3500

CITY: DENVER

STATE: CO

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,460A

FILING DATE: 14-OCT-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 14-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: KOVARIK, JOSEPH E.

REGISTRATION NUMBER: 33,005

REFERENCE/DOCKET NUMBER: 2879-1-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/863-9700

TELEFAX: 303/863-0223

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-323-460A-6

Query Match 90.3%; Score 28; DB 2; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

Db 64 VRYEDV 69

RESULT 15
US-08-461-146C-6
Sequence 6, Application US/08461146C

Patent No. 5981265

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,146C

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/354,516

FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/323,460

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11690

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04178

FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: KARA, Catherine J.

REGISTRATION NUMBER: P41,106

REFERENCE/DOCKET NUMBER: CPI-004CN3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-461-146C-6

Query Match 90.3%; Score 28; DB 2; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

Db 64 VRYEDV 69

RESULT 16
US-08-461-145C-6
Sequence 6, Application US/08461145C

Patent No. 6074861

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: NOVEL MEKK PROTEINS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,145C

FILING DATE: 5-JUNE-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 11-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/354,516
FILING DATE: 21-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P-41,106
REFERENCE/DOCKET NUMBER: CPT-004CNI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-145C-6

Query Match 90.3%; Score 28; DB 3; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||:
DB 64 VRYEDV 69

RESULT 17
US-09-423-890-6
Sequence 6, Application US/09423890
Patent No. 6312934
GENERAL INFORMATION:
APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USSN 60/078,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USSN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 6
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapiens
US-09-423-890-6

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||:
DB 64 VRYEDV 69

RESULT 18
US-09-423-890-12
Sequence 12, Application US/09423890
Patent No. 6312934
GENERAL INFORMATION:

APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USSN 60/078,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USSN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-09-423-890-12

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||:
DB 64 VRYEDV 69

RESULT 19
US-08-628-829-10
Sequence 10, Application US/08628829A
Patent No. 6333170
GENERAL INFORMATION:
APPLICANT: Johnson, Gary L.
TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extr
FILE REFERENCE: CPT-004DVC3
CURRENT APPLICATION NUMBER: US/08/628,829A
PRIOR FILING DATE: 1995-04-05
PRIOR APPLICATION NUMBER: 08/440,421
PRIOR FILING DATE: 1995-05-15
PRIOR APPLICATION NUMBER: 08/323,460
PRIOR FILING DATE: 1994-10-14
PRIOR APPLICATION NUMBER: 08/049,254
PRIOR FILING DATE: 1993-05-15
PRIOR APPLICATION NUMBER: 08/410,602
PRIOR FILING DATE: 1995-04-24
PRIOR APPLICATION NUMBER: 08/472,934
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 10
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-08-628-829-10

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||:
DB 64 VRYEDV 69

RESULT 20
US-08-411-607A-2
Sequence 2, Application US/08411607A
Patent No. 594102
GENERAL INFORMATION:
APPLICANT: HUDSON, PETER L
APPLICANT: ROSEN, CRAIG A
APPLICANT: HE, WEI WU
TITLE OF INVENTION: PROSTATIC GROWTH FACTOR

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
ADDRESSEE: STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NJ
COUNTRY: US
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/411,607A
FILING DATE: 11-Apr-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-329
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1740
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 295 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-411-607A-2

Query Match 87.1%; Score 27; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 46 RYEDL 50

RESULT 21
US-08-927-433-9
Sequence 9, Application US/08927433
Patent No. 6107476
GENERAL INFORMATION:
APPLICANT: Eriander, Mark G.
APPLICANT: Huang, Shaoming
APPLICANT: Jackson, Michael A.
APPLICANT: Peterson, Per A.
TITLE OF INVENTION: PROTEIN GROWTH FACTOR FOR TREATING PROSTATE CANCER, AND RELATE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Johnson & Johnson
STREET: One J & J Plaza
CITY: New Brunswick
STATE: New Jersey
COUNTRY: USA
ZIP: 08933
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,433
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Morrison, Alan J.
REGISTRATION NUMBER: 37,399
REFERENCE/DOCKET NUMBER: CNT-849

TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-524-3592
TELEFAX: 732-524-2808
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-927-433-9

Query Match 87.1%; Score 27; DB 3; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 22
US-08-775-882-4
Sequence 4, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: Oe, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAKI, Mihoro
APPLICANT: KOBAYASHI, Midori
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-4

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 23
US-08-775-882-6
Sequence 6, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: OH, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAKETI, Mihoro
APPLICANT: KOBAYASHI, Midori
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
TITLE OF INVENTION: AGENT
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-6

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 24
US-09-276-600-6
Sequence 6, Application US/09276600
Patent No. 6465181
GENERAL INFORMATION:
APPLICANT: Patricia Billing-Wedel
APPLICANT: Maurice Cohen
APPLICANT: Tracey L. Colpitts
APPLICANT: Julian Gordon
APPLICANT: Edward N. Granados
APPLICANT: John C. Russell
APPLICANT: Stephen D. Stroupe
TITLE OF INVENTION: Reagents and Methods Useful for
TITLE OF INVENTION: Detecting Disease of the Prostate
FILE REFERENCE: 6397.US.01
CURRENT APPLICATION NUMBER: US/09/276,600
CURRENT FILING DATE: 1999-03-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 308
TYPE: PRT
ORGANISM: Homo sapiens
US-09-276-600-6

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 25
US-09-067-351-2
Sequence 2, Application US/09067351
Patent No. 5994081
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,351
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 2029060
US-09-067-351-2

Query Match 87.1%; Score 27; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 237 VRYED 241

RESULT 26
US-09-360-490-2
Sequence 2, Application US/09360490
Patent No. 6221843
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/360,490
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/067,351
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 2029060
US-09-360-490-2

Query Match 87.1%; Score 27; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 237 VRYED 241

RESULT 27
US-09-041-236-4
Sequence 4, Application US/09041236
Patent No. 6225285
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
APPLICANT: Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,236
FILING DATE: March 11, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-236-4

Query Match 87.1%; Score 27; DB 4; Length 606;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 415 VRYED 419

RESULT 28
US-09-412-554A-2
Sequence 2, Application US/09412554A
Patent No. 6355788

GENERAL INFORMATION:
APPLICANT: Conklin, Darrell
APPLICANT: Ellsworth, Jeff L.
TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTX2
FILE REFERENCE: 98-50
CURRENT APPLICATION NUMBER: US/09/412,554A
CURRENT FILING DATE: 1999-10-05
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 983
TYPE: PRN
ORGANISM: Homo sapiens
US-09-412-554A-2

Query Match 87.1%; Score 27; DB 4; Length 983;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYED 5
DB 539 VYED 543

RESULT 29
US-09-396-651B-2
Sequence 2, Application US/09396651B
Patent No. 6225076
GENERAL INFORMATION:
APPLICANT: Darst, Seth A
APPLICANT: Zhang, Gongyi
APPLICANT: Campbell, Elizabeth
APPLICANT: Minakin, Leonid
APPLICANT: Severinov, Konstantin
TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
FILE REFERENCE: 600-1-258
CURRENT APPLICATION NUMBER: US/09/396,651B
CURRENT FILING DATE: 1999-09-15
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1119
TYPE: PRN
FEATURE:
ORGANISM: Thermus aquaticus
NAME/KEY: SITE
LOCATION: (695)..(696)
OTHER INFORMATION: Any amino acids can be at these two positions.
US-09-396-651B-2

Query Match 87.1%; Score 27; DB 4; Length 1119;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYED 5
DB 613 VYED 617

RESULT 30
US-08-117-083-60
Sequence 60, Application US/08117083
Patent No. 5719054
GENERAL INFORMATION:
APPLICANT: Bourisnelli, Michael E.
APPLICANT: Inglis, Stephen C.
APPLICANT: Munro, Alan J.
TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
TITLE OF INVENTION: Papilloma Virus Proteins
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Walter H. Dreger

STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,083
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-58783

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 416 amino acids
TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..416

OTHER INFORMATION:

OTHER INFORMATION: the open reading frame."

US-08-117-083-60

Query Match 83.9%; Score 26; DB 1; Length 416;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYED 6
DB 341 VYED 346

RESULT 31
US-08-588-983-20
Sequence 20, Application US/08588983
Patent No. 5854067

GENERAL INFORMATION:

APPLICANT: Christopher B. Newgard, et al.

TITLE OF INVENTION: Methods and Compositions

TITLE OF INVENTION: for Inhibiting Hexokinase

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P O Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/588,983

FILING DATE: Concurrently herewith

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Fusssey, Shelley P M.

REGISTRATION NUMBER: 39,458

REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: N/A
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-20

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 154 VRHEDL 159

RESULT 32
US-08-588-983-22
Sequence 22, Application US/08588983
Patent No. 5854067

GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: N/A
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-22

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 154 VRHEDL 159

RESULT 33

US-08-588-976-20
Sequence 20, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: N/A
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-20

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 154 VRHEDL 159

RESULT 34
US-08-588-976-22
Sequence 22, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-22

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 154 VRHEDL 159

RESULT 35
US-08-588-983-9
Sequence 9, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-9

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||

Db 608 VRHEDL 613

RESULT 36
US-08-588-983-12
Sequence 12, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-12

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 608 VRHEDL 613

RESULT 37
US-08-588-976-9
Sequence 9, Application US/08588976
Patent No. 5851717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-9

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 608 VRHEDL 613

RESULT 38
US-08-588-976-12
Sequence 12, Application US/08588976
Patent No. 5691717

GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-12

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 608 VRHEDL 613

RESULT 39
US-08-040-751-3
Sequence 3, Application US/08040751
Patent No. 5407825

GENERAL INFORMATION:
APPLICANT: Payne, Jewel M
APPLICANT: Sick, August J
TITLE OF INVENTION: No. 5407825e1 Bacillus thuringiensis isolates
TITLE OF INVENTION: active against lepidopteran pests and genes Encoding No. 54
TITLE OF INVENTION: Lepidopteran-active Toxins
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DAVID R. SALIMANCHIK
STREET: 2421 N.W. 41st STREET, SUITE A-1
CITY: GAINESVILLE
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,751
FILING DATE: 19930329
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SALIMANCHIK, DAVID R.
REFERENCE/DOCKET NUMBER: MA39.C1.D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
MOLECULE TYPE: protein
TOPOLOGY: linear
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: LambdaGem - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-040-751-3

Query Match 83.9%; Score 26; DB 1; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
DB 1087 IRYED 1091

RESULT 40
US-08-291-368-2
Sequence 2, Application US/08291368
Patent No. 5686069
GENERAL INFORMATION:

APPLICANT: Payne, Jewel M.
APPLICANT: Sick, August J.
TITLE OF INVENTION: No. 5686069e1 Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,368
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA50.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: LambdaGen - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-291-368-2
Query Match 83.9%; Score 26; DB 1; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VARIED 5
Db 1087 IRYED 1091
RESULT 41
US-08-962-190-2
Sequence 2, Application US/08962190
Patent No. 5985267
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Sick, August J.
TITLE OF INVENTION: No. 5985267e1 Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/962,190
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/291,368
FILING DATE:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA50.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: LambdaGen - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-962-190-2
Query Match 83.9%; Score 26; DB 2; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VARIED 5
Db 1087 IRYED 1091
RESULT 42
PCT-US95-10310-2
Sequence 2, Application PC/TUS9510310
GENERAL INFORMATION:
APPLICANT: MYCOGEN CORPORATION
APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
APPLICANT: CITY: San Diego
APPLICANT: STATE/PROVINCE: California
APPLICANT: COUNTRY: US
APPLICANT: POSTAL CODE/ZIP: 92121
APPLICANT: PHONE NUMBER: (619) 453-8030
APPLICANT: FAX NUMBER: (619) 453-6991
TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10310
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanhik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA50.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: lambdagem - 11 (tm) Library of August Sick
CLONE: 81A2
PCT-US95-10310-2

Query Match
Best Local Similarity 83.9%; Score 26; DB 5; Length 1174;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 5
Db 1087 IRYED 1091

RESULT 43
5164180-4
; Patent No. 5164180
; APPLICANT: Payne, Jewel; Sick, August J.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS ISOLATES ACTIVE
; AGAINST LEPIDOPTERAN PESTS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/451,389
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 353,860
; FILING DATE: 18-MAY-1989
; SEQ ID NO: 4:
; LENGTH: 1174
5164180-4

Query Match
Best Local Similarity 83.9%; Score 26; DB 6; Length 1174;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 5
Db 1087 IRYED 1091

RESULT 44
```

```

US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/172,422A
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-172-422-1

Query Match
Best Local Similarity 83.9%; Score 26; DB 4; Length 2548;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 6
Db 1920 IRYKDL 1925

RESULT 45
US-09-370-253-14
; Sequence 14, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-370-253-14

Query Match
Best Local Similarity 80.6%; Score 25; DB 4; Length 103;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VYED 6
Db 86 IRYIDL 91

RESULT 46
US-09-352-990-20
; Sequence 20, Application US/09352990
; Patent No. 6255090
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
```

;; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
;; FILE REFERENCE: BB-1191
;; CURRENT APPLICATION NUMBER: US/09/352,990
;; CURRENT FILING DATE: 1999-07-14
;; EARLIER APPLICATION NUMBER: 60/092,866
;; EARLIER FILING DATE: July 15, 1998
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 20
;; LENGTH: 171
;; TYPE: PRT
;; ORGANISM: Trilicium aestivum
US-09-352-990-20

Query Match 80.6%; Score 25; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 128 VRYEEL 133

RESULT 47
US-08-698-805-4
; Sequence 4, Application US/08698805
; Patent No. 5869288
; GENERAL INFORMATION:
; APPLICANT: Chapman, Martin
; APPLICANT: Attitude, L. Karla
; TITLE OF INVENTION: Molecular Cloning of Cockroach
; TITLE OF INVENTION: Allergens, Amino Acid and Nucleotide Sequences Therefore,
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oblon, Spivak, McClelland, Maiter & Neustadt,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/698,805
; FILING DATE: 16-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,510
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 49-203-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-698-805-4

Query Match 80.6%; Score 25; DB 2; Length 182;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 172 LYVEDL 177

RESULT 48
US-09-210-748A-2
; Sequence 2, Application US/09210748A
; Patent No. 6335156
; GENERAL INFORMATION:
; APPLICANT: Hermeking, Helko
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: 14-3-3 SIGMA ARREST THE CELL CYCLE
; FILE REFERENCE: 1107.77810
; CURRENT APPLICATION NUMBER: US/09/210,748A
; CURRENT FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/069,416
; PRIOR FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-210-748A-2

Query Match 80.6%; Score 25; DB 4; Length 248;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 18 RYEDM 22

RESULT 49
US-07-887-072B-2
; Sequence 2, Application US/07887072B
; Patent No. 5424191
; GENERAL INFORMATION:
; APPLICANT: Prasad Ph.D., Gaddamannu L.
; APPLICANT: Cooper M.D., Herbert L.
; TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 W. Broadway, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/887,072B
; FILING DATE: 20-MAY-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kirkpatrick Ph.D., Anita M.
; REGISTRATION NUMBER: 32,617
; REFERENCE/DOCKET NUMBER: NIH021,021A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 amino acids
; TYPE: amino acid

Search completed: February 20, 2003, 13:35:10
Job time : 16.8571 secs

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-07-887-072B-2

Query Match 80.6%; Score 25; DB 1; Length 249;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
Db 18 RYEDM 22

RESULT 50
US-08-466-444-2
Sequence 2, Application US/08466444
Patent No. 5776676
GENERAL INFORMATION:
APPLICANT: Prasad Ph.D., Gaddamanugu L.
APPLICANT: Cooper M.D., Heider L.
TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
MARKER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 W. Broadway, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,444
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/887,072
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kirpatrick Ph.D., Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH021,021A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-466-444-2

Query Match 80.6%; Score 25; DB 1; Length 249;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
Db 18 RYEDM 22

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:33:52 ; Search time 2.74286 Seconds
(without alignments)
55.888 Million cell updates/sec

Title: US-09-816-825-9
Perfect score: 31
Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548976 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	6	10	US-09-816-825-9
2	31	100.0	169	10	US-09-927-602-9
3	31	100.0	171	10	US-09-927-602-8
4	31	100.0	179	10	US-09-927-602-10
5	31	100.0	386	9	US-10-007-262-1
6	31	100.0	386	10	US-09-816-825-2
7	31	100.0	418	9	US-09-927-602-5
8	31	100.0	418	9	US-10-212-933-2
9	31	100.0	484	9	US-10-212-933-4
10	31	100.0	531	10	US-09-833-790-255
11	28	90.3	625	12	US-09-927-602-11
12	28	90.3	625	12	US-10-000-864-6
13	28	90.3	625	12	US-10-000-864-12
14	27	87.1	65	10	US-09-864-761-35570
15	27	87.1	65	10	US-09-864-761-49114
16	27	87.1	157	9	US-09-738-626-3600
17	27	87.1	169	10	US-09-927-602-6
18	27	87.1	169	10	US-09-927-602-7
19	27	87.1	295	10	US-09-361-741-2

20	27	87.1	308	9	US-09-981-353-188	Sequence 188, App
21	27	87.1	308	10	US-09-276-600-6	Sequence 6, Appl1
22	27	87.1	316	9	US-09-966-546-14	Sequence 14, Appl
23	27	87.1	316	9	US-09-966-545-14	Sequence 14, Appl
24	27	87.1	316	9	US-09-965-212-14	Sequence 14, Appl
25	27	87.1	390	10	US-09-927-602-4	Sequence 4, Appl1
26	27	87.1	395	10	US-09-927-602-2	Sequence 2, Appl1
27	27	87.1	395	10	US-09-927-602-3	Sequence 3, Appl1
28	27	87.1	396	10	US-09-815-242-12541	Sequence 12541, A
29	27	87.1	396	10	US-09-815-242-12900	Sequence 12900, A
30	27	87.1	439	10	US-09-741-669-300	Sequence 300, App
31	27	87.1	439	10	US-09-815-242-10284	Sequence 10284, A
32	27	87.1	604	10	US-09-907-509-2	Sequence 2, Appl1
33	27	87.1	606	9	US-09-771-467C-4	Sequence 59, Appl
34	27	87.1	693	9	US-09-808-602-59	Sequence 60, Appl
35	27	87.1	773	9	US-09-808-602-60	Sequence 60, Appl
36	27	87.1	815	9	US-09-808-602-4	Sequence 4, Appl1
37	27	87.1	815	9	US-09-966-546-18	Sequence 18, Appl
38	27	87.1	815	9	US-09-966-545-18	Sequence 18, Appl
39	27	87.1	815	9	US-09-965-212-18	Sequence 18, Appl
40	27	87.1	842	9	US-09-966-546-16	Sequence 16, Appl
41	27	87.1	842	9	US-09-966-545-16	Sequence 16, Appl
42	27	87.1	842	9	US-09-965-212-16	Sequence 16, Appl
43	27	87.1	850	9	US-09-808-602-58	Sequence 58, Appl
44	27	87.1	850	9	US-09-808-602-73	Sequence 73, Appl
45	27	87.1	983	9	US-10-013-136-2	Sequence 2, Appl1
46	27	87.1	1119	10	US-09-782-714-2	Sequence 2, Appl1
47	26	83.9	45	10	US-09-864-761-34609	Sequence 34609, A
48	26	83.9	144	10	US-09-764-878-127	Sequence 127, App
49	26	83.9	144	10	US-09-764-860-161	Sequence 461, App
50	26	83.9	135	9	US-09-764-868-919	Sequence 919, App
51	26	83.9	332	9	US-09-764-868-1225	Sequence 1225, App
52	26	83.9	407	9	US-09-738-626-5541	Sequence 5541, Ap
53	26	83.9	904	9	US-09-964-899-49	Sequence 49, Appl
54	26	83.9	2548	10	US-09-851-682A-1	Sequence 1, Appl1
55	25	80.6	55	10	US-09-925-299-11550	Sequence 1150, App
56	25	80.6	146	10	US-09-747-835A-16	Sequence 16, Appl
57	25	80.6	195	10	US-09-815-028-8	Sequence 8, Appl1
58	25	80.6	248	10	US-09-939-581A-2	Sequence 2, Appl1
59	25	80.6	371	10	US-09-815-242-10916	Sequence 10916, A
60	25	80.6	373	9	US-09-738-626-3568	Sequence 3568, Ap
61	25	80.6	383	10	US-09-771-161A-110	Sequence 110, App
62	25	80.6	414	10	US-09-771-161A-111	Sequence 111, App
63	25	80.6	488	10	US-09-815-242-11444	Sequence 11444, A
64	25	80.6	498	9	US-09-893-519A-33	Sequence 33, Appl
65	25	80.6	599	10	US-09-789-561-120	Sequence 120, App
66	25	80.6	626	10	US-09-771-161A-201	Sequence 201, App
67	25	80.6	632	10	US-09-771-161A-202	Sequence 202, App
68	25	80.6	832	10	US-09-919-585-21	Sequence 21, Appl
69	25	80.6	928	10	US-09-801-368-282	Sequence 282, App
70	24	77.4	30	10	US-09-864-761-45818	Sequence 45818, A
71	24	77.4	67	10	US-09-864-761-39410	Sequence 39410, A
72	24	77.4	79	10	US-09-864-761-45955	Sequence 45955, A
73	24	77.4	93	10	US-09-840-459-466	Sequence 466, App
74	24	77.4	98	10	US-09-840-459-466	Sequence 466, App
75	24	77.4	139	9	US-09-796-692-1084	Sequence 1084, Ap
76	24	77.4	139	9	US-09-796-692-1558	Sequence 1558, Ap
77	24	77.4	139	9	US-09-796-692-1955	Sequence 1955, Ap
78	24	77.4	139	9	US-09-796-692-2194	Sequence 2194, Ap
79	24	77.4	139	9	US-09-796-692-2194	Sequence 2194, Ap
80	24	77.4	143	9	US-09-934-465-112	Sequence 12, Appl
81	24	77.4	148	10	US-09-934-465-112	Sequence 12, Appl
82	24	77.4	186	10	US-09-925-299-1164	Sequence 1164, Ap
83	24	77.4	187	10	US-09-802-116-2	Sequence 2, Appl1
84	24	77.4	209	10	US-09-925-299-10138	Sequence 1038, Ap
85	24	77.4	233	9	US-09-738-626-5648	Sequence 5648, Ap
86	24	77.4	234	10	US-09-813-229-16	Sequence 16, Appl
87	24	77.4	234	10	US-09-921-667-12	Sequence 2, Appl1
88	24	77.4	239	10	US-09-795-440A-2	Sequence 2, Appl1
89	24	77.4	239	10	US-09-795-440A-4	Sequence 4, Appl1
90	24	77.4	239	12	US-10-060-857-1	Sequence 1, Appl1
91	24	77.4	240	10	US-09-819-817-7	Sequence 7, Appl1
92	24	77.4	271	10	US-09-872-733-7	Sequence 7, Appl1
			271	10	US-09-872-733-17	Sequence 17, Appl1

93	24	77.4	271	10	US-09-991-258-7	Sequence 7, Appl1
94	24	77.4	280	9	US-09-738-626-6859	Sequence 6859, Ap
95	24	77.4	280	10	US-09-983-306-1	Sequence 1, Appl1
96	24	77.4	290	10	US-09-729-674-186	Sequence 186, App
97	24	77.4	296	9	US-09-738-626-4271	Sequence 4271, Ap
98	24	77.4	305	10	US-09-815-242-5706	Sequence 5706, Ap
99	24	77.4	309	10	US-09-815-242-11290	Sequence 11290, A
100	24	77.4	311	10	US-09-815-242-10735	Sequence 10735, A
101	24	77.4	315	10	US-09-983-306-3	Sequence 3, Appl1
102	24	77.4	315	10	US-09-983-306-6	Sequence 6, Appl1
103	24	77.4	347	10	US-09-815-242-12466	Sequence 12466, A
104	24	77.4	359	10	US-09-881-752A-248	Sequence 248, App
105	24	77.4	367	10	US-09-779-427-2	Sequence 2, Appl1
106	24	77.4	391	10	US-09-934-249-7	Sequence 7, Appl1
107	24	77.4	398	10	US-09-925-302-472	Sequence 472, App
108	24	77.4	398	10	US-09-815-242-13377	Sequence 13377, A
109	24	77.4	405	9	US-09-738-626-6717	Sequence 6717, Ap
110	24	77.4	420	10	US-09-946-142-2	Sequence 2, Appl1
111	24	77.4	426	10	US-09-815-242-10569	Sequence 10569, A
112	24	77.4	458	10	US-09-827-040-9	Sequence 9, Appl1
113	24	77.4	482	10	US-09-804-551B-32	Sequence 32, Appl1
114	24	77.4	487	9	US-10-033-245-18	Sequence 18, Appl1
115	24	77.4	487	9	US-10-033-223-18	Sequence 18, Appl1
116	24	77.4	487	9	US-10-033-167-18	Sequence 18, Appl1
117	24	77.4	487	9	US-10-033-244-18	Sequence 18, Appl1
118	24	77.4	487	9	US-10-028-072-528	Sequence 528, App
119	24	77.4	487	9	US-10-121-049-528	Sequence 528, App
120	24	77.4	487	9	US-10-123-904-528	Sequence 528, App
121	24	77.4	487	9	US-10-140-470-528	Sequence 528, App
122	24	77.4	487	9	US-10-033-435-18	Sequence 18, Appl1
123	24	77.4	487	9	US-10-175-746-528	Sequence 528, App
124	24	77.4	487	9	US-10-166-918-528	Sequence 528, App
125	24	77.4	487	9	US-10-176-921-528	Sequence 528, App
126	24	77.4	487	9	US-10-227-884-304	Sequence 204, App
127	24	77.4	487	12	US-10-033-246-18	Sequence 18, Appl1
128	24	77.4	487	12	US-10-033-301-18	Sequence 18, Appl1
129	24	77.4	487	12	US-10-033-326-18	Sequence 18, Appl1
130	24	77.4	490	9	US-09-905-291A-132	Sequence 132, App
131	24	77.4	490	9	US-09-902-853-132	Sequence 132, App
132	24	77.4	490	9	US-09-907-824-132	Sequence 132, App
133	24	77.4	490	9	US-09-907-841-132	Sequence 132, App
134	24	77.4	490	9	US-09-904-011-132	Sequence 132, App
135	24	77.4	490	9	US-10-028-072-310	Sequence 310, App
136	24	77.4	490	9	US-09-966-742-132	Sequence 132, App
137	24	77.4	490	9	US-10-121-049-310	Sequence 310, App
138	24	77.4	490	9	US-10-123-904-310	Sequence 310, App
139	24	77.4	490	9	US-10-140-470-310	Sequence 310, App
140	24	77.4	490	9	US-09-906-838-132	Sequence 132, App
141	24	77.4	490	9	US-09-907-613-132	Sequence 132, App
142	24	77.4	490	9	US-09-907-942-132	Sequence 132, App
143	24	77.4	490	9	US-10-175-746-310	Sequence 310, App
144	24	77.4	490	9	US-10-176-918-310	Sequence 310, App
145	24	77.4	490	9	US-10-176-921-310	Sequence 310, App
146	24	77.4	490	10	US-09-909-320-132	Sequence 132, App
147	24	77.4	490	10	US-09-909-088B-132	Sequence 132, App
148	24	77.4	514	10	US-09-800-729-124	Sequence 124, App
149	24	77.4	515	10	US-09-970-711-8	Sequence 8, Appl1
150	24	77.4	520	9	US-09-738-626-4462	Sequence 4462, Ap
151	24	77.4	520	10	US-09-815-242-5735	Sequence 5735, Ap
152	24	77.4	521	10	US-09-848-035-6	Sequence 6, Appl1
153	24	77.4	521	10	US-09-986-224-6	Sequence 6, Appl1
154	24	77.4	562	10	US-09-815-242-5003	Sequence 5003, Ap
155	24	77.4	593	10	US-09-815-242-10733	Sequence 10733, A
156	24	77.4	593	10	US-09-802-741A-1	Sequence 1, Appl1
157	24	77.4	608	10	US-09-764-864-1121	Sequence 1121, Ap
158	24	77.4	609	9	US-10-003-392-4	Sequence 4, Appl1
159	24	77.4	613	10	US-09-971-309-5	Sequence 5, Appl1
160	24	77.4	615	9	US-10-003-392-17	Sequence 17, Appl1
161	24	77.4	693	10	US-09-741-669-393	Sequence 393, App
162	24	77.4	724	9	US-09-925-302-472	Sequence 2, Appl1
163	24	77.4	732	10	US-09-991-496-18	Sequence 18, Appl1
164	24	77.4	732	10	US-09-759-010-5	Sequence 5, Appl1
165	24	77.4	732	10	US-09-874-923-18	Sequence 18, Appl1
166	24	77.4	792	9	US-10-025-380-1127	Sequence 1127, Ap
167	24	77.4	913	10	US-09-815-242-5218	Sequence 5218, Ap
168	24	77.4	946	10	US-09-828-423-3	Sequence 3, Appl1
169	24	77.4	1016	10	US-09-986-224-19	Sequence 19, Appl1
170	24	77.4	1256	9	US-10-047-542-90	Sequence 90, Appl1
171	24	77.4	1265	9	US-09-964-899-19	Sequence 19, Appl1
172	24	77.4	1352	9	US-09-736-968A-9	Sequence 9, Appl1
173	24	77.4	1352	9	US-09-736-968A-9	Sequence 9, Appl1
174	24	77.4	1352	10	US-09-736-968A-9	Sequence 9, Appl1
175	24	77.4	1352	10	US-09-736-968A-9	Sequence 9, Appl1
176	24	77.4	1354	9	US-09-736-968A-10	Sequence 10, Appl1
177	24	77.4	1354	9	US-09-736-968A-10	Sequence 10, Appl1
178	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
179	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
180	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
181	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
182	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
183	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
184	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
185	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
186	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
187	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
188	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
189	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
190	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
191	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
192	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
193	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
194	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
195	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
196	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
197	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
198	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
199	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
200	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
201	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
202	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
203	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
204	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
205	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
206	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
207	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
208	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
209	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
210	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
211	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
212	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
213	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
214	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
215	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
216	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
217	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
218	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
219	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
220	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
221	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
222	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
223	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
224	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
225	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
226	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
227	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
228	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
229	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
230	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
231	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
232	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
233	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
234	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
235	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
236	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
237	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1
238	24	77.4	1354	10	US-09-736-968A-10	Sequence 10, Appl1

239	23	74.2	313	10	US-09-815-242-10920	Sequence 10920, A	312	23	74.2	858	10	US-09-986-224-8	Sequence 8, Appl1
240	23	74.2	345	10	US-09-815-242-13259	Sequence 13259, A	313	23	74.2	876	10	US-09-815-242-13003	Sequence 13003, A
241	23	74.2	357	10	US-09-815-242-10826	Sequence 10826, A	314	23	74.2	879	10	US-09-815-242-5312	Sequence 5312, Ap
242	23	74.2	357	10	US-09-815-242-11389	Sequence 11389, A	315	23	74.2	883	10	US-09-815-242-13382	Sequence 13382, A
243	23	74.2	359	10	US-09-815-242-13336	Sequence 13336, A	316	23	74.2	883	10	US-09-815-242-13364	Sequence 13364, A
244	23	74.2	359	10	US-09-799-777-45	Sequence 45, Appl	317	23	74.2	883	10	US-09-815-242-13838	Sequence 13838, A
245	23	74.2	360	10	US-09-815-242-5195	Sequence 5195, Ap	318	23	74.2	900	9	US-09-738-626-6108	Sequence 6108, Ap
246	23	74.2	361	10	US-09-815-242-11251	Sequence 11251, A	319	23	74.2	904	10	US-09-894-998-18	Sequence 18, Appl
247	23	74.2	371	10	US-09-828-302-15	Sequence 15, Appl	320	23	74.2	915	10	US-09-817-514A-6	Sequence 6, Appl
248	23	74.2	373	9	US-10-137-731-2	Sequence 2, Appl1	321	23	74.2	951	10	US-09-815-242-10465	Sequence 10465, A
249	23	74.2	380	9	US-09-738-626-4866	Sequence 4866, Ap	322	23	74.2	951	10	US-09-815-242-13853	Sequence 13853, A
250	23	74.2	392	10	US-09-764-848-36	Sequence 36, Appl	323	23	74.2	994	9	US-10-066-571-24	Sequence 24, Appl
251	23	74.2	396	9	US-09-971-536-64	Sequence 64, Appl	324	23	74.2	1037	10	US-09-728-721-55	Sequence 55, Appl
252	23	74.2	415	10	US-09-971-361-12	Sequence 12, Appl	325	23	74.2	1127	10	US-09-815-242-5373	Sequence 5373, Ap
253	23	74.2	422	10	US-09-779-307-18	Sequence 18, Appl	326	23	74.2	1148	10	US-09-815-242-11895	Sequence 11895, A
254	23	74.2	422	10	US-09-779-307-20	Sequence 20, Appl	327	23	74.2	1153	10	US-09-992-056-2	Sequence 2, Appl1
255	23	74.2	423	9	US-09-738-626-3778	Sequence 3778, Ap	328	23	74.2	1158	10	US-09-815-242-12522	Sequence 12522, A
256	23	74.2	424	10	US-09-733-524-16	Sequence 16, Appl	329	23	74.2	1242	10	US-09-904-065-12	Sequence 12, Appl
257	23	74.2	425	9	US-10-120-319-6	Sequence 6, Appl1	330	23	74.2	1242	10	US-09-904-065-13	Sequence 13, Appl
258	23	74.2	428	10	US-09-779-307-21	Sequence 21, Appl	331	23	74.2	1332	9	US-10-041-865-3	Sequence 3, Appl1
259	23	74.2	430	10	US-09-779-307-19	Sequence 19, Appl	332	23	74.2	1333	9	US-09-738-626-4048	Sequence 4048, Ap
260	23	74.2	432	10	US-09-815-242-5096	Sequence 2, Appl1	333	23	74.2	1706	10	US-09-864-761-46862	Sequence 46862, A
261	23	74.2	454	9	US-10-060-425-17	Sequence 17, Appl	334	23	74.2	1797	10	US-09-881-752A-308	Sequence 308, App
262	23	74.2	454	9	US-10-120-319-8	Sequence 8, Appl1	335	23	74.2	1958	12	US-10-028-446-4	Sequence 4, Appl1
263	23	74.2	454	10	US-09-827-040-4	Sequence 4, Appl1	336	23	74.2	2053	9	US-10-017-216-2	Sequence 2, Appl1
264	23	74.2	454	10	US-09-733-524-18	Sequence 18, Appl	337	23	74.2	2054	12	US-10-028-446-2	Sequence 2, Appl1
265	23	74.2	456	10	US-09-815-242-10870	Sequence 10870, A	338	23	74.2	2055	9	US-10-017-216-4	Sequence 4, Appl1
266	23	74.2	464	9	US-10-120-319-1	Sequence 1, Appl1	339	23	74.2	2890	10	US-09-815-242-11410	Sequence 11410, A
267	23	74.2	473	10	US-09-815-242-11235	Sequence 11235, A	340	23	74.2	2890	10	US-09-815-242-11571	Sequence 11571, A
268	23	74.2	476	9	US-10-120-319-5	Sequence 5, Appl1	341	22	71.0	26	10	US-09-870-759-128	Sequence 128, App
269	23	74.2	476	10	US-09-733-524-15	Sequence 15, Appl	342	22	71.0	26	10	US-09-749-334A-5	Sequence 5, Appl1
270	23	74.2	476	10	US-09-779-307-16	Sequence 16, Appl	343	22	71.0	41	10	US-09-864-761-37339	Sequence 37339, A
271	23	74.2	477	10	US-09-779-307-4	Sequence 4, Appl1	344	22	71.0	46	10	US-09-864-761-42065	Sequence 42065, A
272	23	74.2	478	9	US-10-120-319-7	Sequence 7, Appl1	345	22	71.0	46	10	US-09-764-877-11937	Sequence 11937, Ap
273	23	74.2	479	10	US-09-733-524-17	Sequence 17, Appl	346	22	71.0	50	10	US-09-938-315-50	Sequence 50, Appl
274	23	74.2	480	10	US-09-802-472B-4	Sequence 4, Appl1	347	22	71.0	52	10	US-09-864-761-42113	Sequence 42113, A
275	23	74.2	482	10	US-09-779-307-17	Sequence 17, Appl	348	22	71.0	62	9	US-09-796-632-1724	Sequence 1724, Ap
276	23	74.2	483	9	US-09-974-298-41	Sequence 41, Appl	349	22	71.0	62	10	US-09-864-761-39964	Sequence 39964, A
277	23	74.2	483	10	US-09-981-353-158	Sequence 158, App	350	22	71.0	70	10	US-09-925-999-12218	Sequence 12218, Ap
278	23	74.2	483	10	US-09-919-497-79	Sequence 79, Appl	351	22	71.0	78	9	US-09-940-240-8	Sequence 8, Appl1
279	23	74.2	485	9	US-09-733-524-2	Sequence 2, Appl1	352	22	71.0	75	9	US-09-738-626-4440	Sequence 4440, Ap
280	23	74.2	486	10	US-10-120-319-2	Sequence 2, Appl1	353	22	71.0	89	9	US-09-738-626-5666	Sequence 5666, Ap
281	23	74.2	490	9	US-09-895-913A-6	Sequence 6, Appl1	354	22	71.0	93	10	US-09-764-860-518	Sequence 518, App
282	23	74.2	492	9	US-09-796-753-42	Sequence 42, Appl1	355	22	71.0	94	9	US-09-940-240-21	Sequence 21, Appl
283	23	74.2	494	9	US-10-238-709-4	Sequence 4, Appl1	356	22	71.0	99	9	US-09-940-240-2	Sequence 2, Appl1
284	23	74.2	494	10	US-09-804-471A-4	Sequence 4, Appl1	357	22	71.0	99	9	US-09-978-280A-5	Sequence 5, Appl1
285	23	74.2	495	10	US-09-815-242-10315	Sequence 10315, A	358	22	71.0	111	9	US-09-978-687-370	Sequence 370, App
286	23	74.2	497	10	US-10-238-709-2	Sequence 2, Appl1	359	22	71.0	111	9	US-09-978-182A-370	Sequence 370, App
287	23	74.2	497	10	US-09-804-471A-2	Sequence 2, Appl1	360	22	71.0	111	9	US-09-999-832A-370	Sequence 370, App
288	23	74.2	501	9	US-09-738-626-6834	Sequence 6834, Ap	361	22	71.0	111	9	US-09-978-189-370	Sequence 370, App
289	23	74.2	501	10	US-09-733-524-1	Sequence 1, Appl1	362	22	71.0	111	9	US-09-816-920-2	Sequence 2, Appl1
290	23	74.2	502	10	US-09-895-072-13	Sequence 13, Appl	363	22	71.0	111	12	US-10-105-282-2	Sequence 2, Appl1
291	23	74.2	502	10	US-09-986-552-13	Sequence 13, Appl	364	22	71.0	111	10	US-09-815-242-5286	Sequence 5286, Ap
292	23	74.2	528	9	US-09-738-626-5577	Sequence 5577, Ap	365	22	71.0	117	10	US-10-014-717-12	Sequence 12, Appl
293	23	74.2	551	10	US-09-802-472B-21	Sequence 21, Appl	366	22	71.0	126	9	US-09-852-767-2	Sequence 2, Appl1
294	23	74.2	560	10	US-09-864-761-37938	Sequence 37938, A	367	22	71.0	133	10	US-09-734-569-30	Sequence 30, Appl1
295	23	74.2	572	10	US-09-864-761-44546	Sequence 34546, A	368	22	71.0	144	10	US-09-764-868-1022	Sequence 1022, Ap
296	23	74.2	590	9	US-09-331-631A-8	Sequence 8, Appl1	369	22	71.0	149	9	US-09-804-868-1022	Sequence 8, Appl1
297	23	74.2	607	9	US-09-344-882-18	Sequence 16, Appl	370	22	71.0	151	10	US-09-815-242-12807	Sequence 12807, A
298	23	74.2	619	10	US-09-738-626-5531	Sequence 5531, Ap	371	22	71.0	151	10	US-10-014-717-10	Sequence 10, Appl1
299	23	74.2	729	10	US-09-815-242-5520	Sequence 5520, Ap	372	22	71.0	155	9	US-09-728-626-6468	Sequence 6468, Ap
300	23	74.2	750	10	US-09-815-242-10783	Sequence 10783, A	373	22	71.0	158	9	US-09-925-302-593	Sequence 593, Appl
301	23	74.2	762	10	US-09-391-340-10	Sequence 10, Appl	374	22	71.0	162	10	US-09-738-626-6015	Sequence 6015, App
302	23	74.2	762	10	US-09-948-369-10	Sequence 10, Appl	375	22	71.0	176	10	US-09-815-242-11338	Sequence 11351, A
303	23	74.2	782	10	US-09-815-242-12367	Sequence 12367, A	376	22	71.0	187	10	US-09-917-265-102	Sequence 102, App
304	23	74.2	782	10	US-09-815-242-12984	Sequence 12984, A	377	22	71.0	187	10	US-09-925-297-490	Sequence 490, App
305	23	74.2	788	10	US-09-978-248-2	Sequence 2, Appl1	378	22	71.0	193	10	US-09-738-626-3784	Sequence 3784, Appl
306	23	74.2	790	10	US-09-765-519-2	Sequence 2, Appl1	379	22	71.0	194	9	US-09-941-936A-2	Sequence 2, Appl1
307	23	74.2	810	10	US-09-815-242-12636	Sequence 12636, A	380	22	71.0	194	9	US-10-025-380-1126	Sequence 1126, Ap
308	23	74.2	820	10	US-09-815-242-5882	Sequence 5882, Ap	381	22	71.0	194	9	US-09-738-626-6465	Sequence 6465, Ap
309	23	74.2	825	10	US-09-815-242-13079	Sequence 13079, A	382	22	71.0				
310	23	74.2	844	10	US-09-815-242-10039	Sequence 10039, A	383	22	71.0				
311	23	74.2	858	10	US-09-848-035-8	Sequence 8, Appl1	384	22	71.0				

385	22	71.0	195	9	US-10-074-956-23	Sequence 23, Appl	458	22	71.0	264	9	US-09-944-160-49	Sequence 49, Appl
386	22	71.0	197	10	US-09-917-265-36	Sequence 36, Appl	459	22	71.0	266	9	US-09-738-626-4442	Sequence 4442, Ap
387	22	71.0	197	10	US-09-917-265-50	Sequence 50, Appl	460	22	71.0	266	10	US-09-815-242-13460	Sequence 13460, A
388	22	71.0	197	10	US-10-079-616-4	Sequence 4, Appl	461	22	71.0	266	10	US-09-815-242-13665	Sequence 13665, A
389	22	71.0	202	12	US-09-925-299-974	Sequence 974, Ap	462	22	71.0	266	10	US-09-925-197-2	Sequence 2, Appl
390	22	71.0	203	9	US-10-174-590-270	Sequence 270, App	463	22	71.0	267	10	US-09-853-555-2	Sequence 2, Appl
391	22	71.0	203	9	US-10-176-758-270	Sequence 270, App	464	22	71.0	268	10	US-09-925-300-1461	Sequence 241, Ap
392	22	71.0	203	9	US-10-175-737-270	Sequence 270, App	465	22	71.0	269	9	US-10-108-605-241	Sequence 4919, Ap
393	22	71.0	203	9	US-10-173-706-270	Sequence 270, App	466	22	71.0	274	10	US-09-815-242-5621	Sequence 5621, Ap
394	22	71.0	203	9	US-10-175-738-270	Sequence 270, App	467	22	71.0	275	9	US-09-974-879-144	Sequence 144, Ap
395	22	71.0	203	9	US-10-175-752-270	Sequence 270, App	468	22	71.0	276	10	US-09-815-242-10555	Sequence 10555, A
396	22	71.0	203	9	US-10-176-482-270	Sequence 270, App	469	22	71.0	282	10	US-09-815-242-11733	Sequence 11733, A
397	22	71.0	203	9	US-10-176-757-270	Sequence 270, App	470	22	71.0	287	10	US-09-734-569-152	Sequence 152, Ap
398	22	71.0	203	9	US-10-176-913-270	Sequence 270, App	471	22	71.0	288	10	US-09-755-016-6	Sequence 6, Appl
399	22	71.0	203	9	US-10-180-552-270	Sequence 270, App	472	22	71.0	290	10	US-09-764-853-661	Sequence 661, Ap
400	22	71.0	203	9	US-10-180-557-270	Sequence 270, App	473	22	71.0	291	10	US-09-870-501-1	Sequence 1, Appl
401	22	71.0	203	9	US-10-173-700-270	Sequence 270, App	474	22	71.0	298	9	US-10-138-634-3	Sequence 3, Appl
402	22	71.0	203	9	US-10-174-572-270	Sequence 270, App	475	22	71.0	298	10	US-09-925-301-1253	Sequence 1253, Ap
403	22	71.0	203	9	US-10-174-579-270	Sequence 270, App	476	22	71.0	300	10	US-09-815-242-11733	Sequence 11733, A
404	22	71.0	203	9	US-10-174-582-270	Sequence 270, App	477	22	71.0	300	10	US-09-815-242-13847	Sequence 13847, A
405	22	71.0	203	9	US-10-174-588-270	Sequence 270, App	478	22	71.0	301	8	US-08-902-572-14	Sequence 14, Appl
406	22	71.0	203	9	US-10-175-739-270	Sequence 270, App	479	22	71.0	301	10	US-09-863-179A-41	Sequence 41, Appl
407	22	71.0	203	9	US-10-175-740-270	Sequence 270, App	480	22	71.0	302	10	US-09-893-737-110	Sequence 310, App
408	22	71.0	203	9	US-10-175-743-270	Sequence 270, App	481	22	71.0	305	10	US-09-828-302-13	Sequence 13, Appl
409	22	71.0	203	9	US-10-176-488-270	Sequence 270, App	482	22	71.0	308	9	US-09-738-626-6317	Sequence 6317, Ap
410	22	71.0	203	9	US-10-176-92-270	Sequence 270, App	483	22	71.0	309	9	US-10-051-643-118	Sequence 118, App
411	22	71.0	203	9	US-10-176-747-270	Sequence 270, App	484	22	71.0	309	9	US-09-880-505-118	Sequence 118, App
412	22	71.0	203	9	US-10-176-750-270	Sequence 270, App	485	22	71.0	310	10	US-09-925-300-1602	Sequence 1602, Ap
413	22	71.0	203	9	US-10-176-985-270	Sequence 270, App	486	22	71.0	316	10	US-09-815-242-13294	Sequence 13294, A
414	22	71.0	203	9	US-10-176-987-270	Sequence 270, App	487	22	71.0	317	10	US-09-815-242-13524	Sequence 13524, A
415	22	71.0	203	9	US-10-176-991-270	Sequence 270, App	488	22	71.0	320	10	US-09-854-132-20	Sequence 20, Appl
416	22	71.0	203	9	US-10-176-993-270	Sequence 270, App	489	22	71.0	322	9	US-10-147-025-2	Sequence 2, Appl
417	22	71.0	203	9	US-10-176-993-270	Sequence 270, App	490	22	71.0	323	10	US-09-815-242-5227	Sequence 5227, Ap
418	22	71.0	203	9	US-10-184-658-270	Sequence 270, App	491	22	71.0	324	10	US-09-815-242-12634	Sequence 12634, A
419	22	71.0	203	12	US-10-052-586-270	Sequence 270, App	492	22	71.0	324	10	US-09-815-242-13004	Sequence 13004, A
420	22	71.0	213	10	US-09-922-958-2	Sequence 2, Appl	493	22	71.0	327	9	US-10-051-643-162	Sequence 162, App
421	22	71.0	213	10	US-09-755-016-2	Sequence 2, Appl	494	22	71.0	327	9	US-09-880-505-162	Sequence 162, App
422	22	71.0	213	10	US-09-754-014-5	Sequence 5, Appl	495	22	71.0	327	9	US-09-765-061B-75	Sequence 75, Appl
423	22	71.0	213	10	US-09-835-865-3	Sequence 3, Appl	496	22	71.0	328	9	US-09-765-061B-76	Sequence 76, Appl
424	22	71.0	213	10	US-09-815-242-10461	Sequence 10461, A	497	22	71.0	328	10	US-09-777-480A-29	Sequence 29, Appl
425	22	71.0	220	10	US-09-815-242-13988	Sequence 13988, A	498	22	71.0	328	10	US-09-777-480A-29	Sequence 29, Appl
426	22	71.0	222	10	US-09-917-265-33	Sequence 33, Appl	499	22	71.0	329	10	US-09-804-969-4	Sequence 4, Appl
427	22	71.0	222	10	US-09-917-265-47	Sequence 47, Appl	500	22	71.0	331	9	US-09-738-626-4960	Sequence 4960, Ap
428	22	71.0	222	10	US-09-917-265-105	Sequence 105, Appl	501	22	71.0	331	10	US-09-738-626-4865	Sequence 4865, App
429	22	71.0	222	12	US-10-079-616-3	Sequence 3, Appl	502	22	71.0	333	10	US-09-939-484-2	Sequence 2, Appl
430	22	71.0	226	9	US-09-976-673-8	Sequence 8, Appl	503	22	71.0	336	10	US-09-939-484-2	Sequence 2, Appl
431	22	71.0	226	9	US-09-976-673-10	Sequence 10, Appl	504	22	71.0	336	10	US-09-939-484-2	Sequence 2, Appl
432	22	71.0	226	9	US-09-976-673-26	Sequence 26, Appl	505	22	71.0	339	9	US-09-981-876-246	Sequence 246, App
433	22	71.0	227	9	US-09-976-673-2	Sequence 2, Appl	506	22	71.0	339	9	US-09-148-543-246	Sequence 246, App
434	22	71.0	227	9	US-09-976-673-4	Sequence 4, Appl	507	22	71.0	342	9	US-09-712-363-210	Sequence 210, App
435	22	71.0	227	9	US-09-976-673-6	Sequence 6, Appl	508	22	71.0	344	9	US-09-810-506-1	Sequence 1, Appl
436	22	71.0	227	9	US-09-976-673-12	Sequence 12, Appl	509	22	71.0	346	10	US-09-925-301-1071	Sequence 1071, Ap
437	22	71.0	227	9	US-09-976-673-14	Sequence 14, Appl	510	22	71.0	348	10	US-09-755-016-4	Sequence 4, Appl
438	22	71.0	227	9	US-09-976-673-24	Sequence 24, Appl	511	22	71.0	349	9	US-10-260-877-10	Sequence 10, Appl
439	22	71.0	229	12	US-10-038-107A-3	Sequence 3, Appl	512	22	71.0	351	10	US-09-804-969-6	Sequence 6, Appl
440	22	71.0	238	9	US-09-870-759-104	Sequence 104, App	513	22	71.0	357	10	US-09-823-414-9	Sequence 9, Appl
441	22	71.0	238	10	US-09-823-356-1	Sequence 1, Appl	514	22	71.0	359	10	US-09-888-615-114	Sequence 114, App
442	22	71.0	239	10	US-09-804-969-2	Sequence 2, Appl	515	22	71.0	361	9	US-09-738-626-3735	Sequence 3735, Ap
443	22	71.0	239	10	US-09-804-969-10	Sequence 10, Appl	516	22	71.0	363	10	US-09-925-300-1530	Sequence 1530, Ap
444	22	71.0	241	9	US-10-074-956-23	Sequence 3674, Ap	517	22	71.0	365	9	US-10-029-180-62	Sequence 62, Appl
445	22	71.0	247	9	US-09-738-626-5674	Sequence 5674, Ap	518	22	71.0	369	12	US-10-062-254-14	Sequence 14, Appl
446	22	71.0	248	10	US-09-853-450-28	Sequence 28, Appl	519	22	71.0	370	10	US-09-815-242-5328	Sequence 5328, Ap
447	22	71.0	248	10	US-09-815-242-12098	Sequence 12098, A	520	22	71.0	370	10	US-09-815-242-13553	Sequence 13553, A
448	22	71.0	250	9	US-09-853-450-30	Sequence 30, Appl	521	22	71.0	372	9	US-10-029-180-10	Sequence 10, Appl
449	22	71.0	250	9	US-09-738-626-9948	Sequence 4948, Ap	522	22	71.0	372	9	US-09-765-061B-74	Sequence 74, Appl
450	22	71.0	253	10	US-09-924-703-6	Sequence 6, Appl	523	22	71.0	372	10	US-09-765-061B-78	Sequence 78, Appl
451	22	71.0	253	12	US-10-042-417-60	Sequence 60, Appl	524	22	71.0	372	10	US-09-766-894A-2	Sequence 2, Appl
452	22	71.0	257	10	US-09-925-299-891	Sequence 891, App	525	22	71.0	374	10	US-09-815-242-13447	Sequence 13447, A
453	22	71.0	258	10	US-09-815-242-11725	Sequence 11725, A	526	22	71.0	374	10	US-09-815-242-130652	Sequence 130652, A
454	22	71.0	261	10	US-09-804-969-12	Sequence 12, Appl	527	22	71.0	375	10	US-09-815-242-13916	Sequence 13916, A
455	22	71.0	264	9	US-09-944-160-47	Sequence 47, Appl	528	22	71.0	377	10	US-09-815-242-13916	Sequence 13916, A
456	22	71.0	264	9	US-09-944-160-45	Sequence 45, Appl	529	22	71.0	383	10	US-09-864-761-48697	Sequence 48697, A
457	22	71.0	264	9	US-09-944-160-48	Sequence 48, Appl	530	22	71.0	384	9	US-09-924-400-333	Sequence 333, App

531	22	71.0	384	9	US-09-924-400-340	Sequence 340, App	604	22	71.0	517	10	US-09-909-903-21	Sequence 21, Appl
532	22	71.0	384	9	US-09-765-061B-72	Sequence 72, Appl	605	22	71.0	518	10	US-09-815-442-12279	Sequence 12279, A
533	22	71.0	384	9	US-09-765-061B-73	Sequence 73, Appl	606	22	71.0	522	10	US-09-765-772-120	Sequence 120, App
534	22	71.0	384	10	US-09-815-242-12644	Sequence 12644, A	607	22	71.0	523	9	US-10-042-141-53	Sequence 53, Appl
535	22	71.0	384	10	US-09-815-242-12883	Sequence 12883, A	608	22	71.0	523	9	US-10-051-663-114	Sequence 114, App
536	22	71.0	384	10	US-09-810-936-333	Sequence 333, App	609	22	71.0	523	9	US-09-880-505-114	Sequence 114, App
537	22	71.0	385	9	US-10-029-180-42	Sequence 42, Appl	610	22	71.0	523	10	US-09-726-643-53	Sequence 53, Appl
538	22	71.0	385	10	US-09-815-242-11483	Sequence 11483, A	611	22	71.0	526	10	US-09-828-825-2	Sequence 2, Appl
539	22	71.0	386	10	US-09-737-178-127	Sequence 127, App	612	22	71.0	527	10	US-09-828-825-2	Sequence 2, Appl
540	22	71.0	387	10	US-09-815-242-10416	Sequence 10416, A	613	22	71.0	528	10	US-09-815-442-5266	Sequence 5266, App
541	22	71.0	387	10	US-09-815-242-13750	Sequence 13750, A	614	22	71.0	533	10	US-09-917-265-39	Sequence 39, Appl
542	22	71.0	390	10	US-09-939-521-10	Sequence 10, Appl	615	22	71.0	533	10	US-09-917-265-62	Sequence 62, Appl
543	22	71.0	391	10	US-09-866-562-62	Sequence 62, Appl	616	22	71.0	537	10	US-09-738-666-4471	Sequence 4471, App
544	22	71.0	392	9	US-09-765-061B-77	Sequence 77, Appl	617	22	71.0	537	10	US-09-828-825-8	Sequence 8, Appl
545	22	71.0	393	10	US-09-815-242-5366	Sequence 5366, App	618	22	71.0	537	10	US-09-828-825-16	Sequence 16, Appl
546	22	71.0	393	10	US-09-784-077-2	Sequence 2, Appl	619	22	71.0	540	9	US-09-713-363-169	Sequence 169, App
547	22	71.0	398	10	US-09-815-242-13171	Sequence 13171, A	620	22	71.0	540	10	US-09-847-637B-6	Sequence 6, Appl
548	22	71.0	411	10	US-09-815-242-12606	Sequence 12606, A	621	22	71.0	541	9	US-10-051-663-160	Sequence 160, App
549	22	71.0	412	9	US-10-029-180-84	Sequence 84, Appl	622	22	71.0	541	9	US-09-880-505-160	Sequence 160, App
550	22	71.0	415	10	US-09-815-242-12099	Sequence 12099, A	623	22	71.0	542	10	US-09-828-825-6	Sequence 6, Appl
551	22	71.0	421	9	US-09-943-123-8	Sequence 8, Appl	624	22	71.0	542	10	US-09-828-825-14	Sequence 14, Appl
552	22	71.0	421	10	US-09-813-355-4	Sequence 4, Appl	625	22	71.0	543	10	US-09-927-738-2	Sequence 2, Appl
553	22	71.0	423	9	US-09-911-317-4	Sequence 4, Appl	626	22	71.0	547	10	US-09-828-825-10	Sequence 10, Appl
554	22	71.0	425	9	US-10-174-590-570	Sequence 570, App	627	22	71.0	547	10	US-09-828-825-12	Sequence 12, Appl
555	22	71.0	425	9	US-10-176-758-570	Sequence 570, App	628	22	71.0	549	9	US-09-959-845-2	Sequence 2, Appl
556	22	71.0	425	9	US-10-175-737-570	Sequence 570, App	629	22	71.0	552	10	US-09-835-654-2	Sequence 2, Appl
557	22	71.0	425	9	US-10-173-706-570	Sequence 570, App	630	22	71.0	555	9	US-10-024-869-2	Sequence 2, Appl
558	22	71.0	425	9	US-10-175-738-570	Sequence 570, App	631	22	71.0	574	9	US-09-990-385-13	Sequence 13, Appl
559	22	71.0	425	9	US-10-175-752-570	Sequence 570, App	632	22	71.0	577	10	US-09-815-442-10193	Sequence 10193, A
560	22	71.0	425	9	US-10-176-482-570	Sequence 570, App	633	22	71.0	577	10	US-09-815-442-11255	Sequence 11255, A
561	22	71.0	425	9	US-10-176-757-570	Sequence 570, App	634	22	71.0	577	10	US-09-815-442-13752	Sequence 13752, A
562	22	71.0	425	9	US-10-176-913-570	Sequence 570, App	635	22	71.0	580	9	US-09-928-457-37	Sequence 37, Appl
563	22	71.0	425	9	US-10-180-552-570	Sequence 570, App	636	22	71.0	582	10	US-09-916-658-4	Sequence 4, Appl
564	22	71.0	425	9	US-10-180-557-570	Sequence 570, App	637	22	71.0	582	10	US-08-801-196-27	Sequence 27, Appl
565	22	71.0	425	9	US-10-173-700-570	Sequence 570, App	638	22	71.0	582	10	US-09-919-997-84	Sequence 84, Appl
566	22	71.0	425	9	US-10-174-572-570	Sequence 570, App	639	22	71.0	583	10	US-09-815-442-10102	Sequence 10102, A
567	22	71.0	425	9	US-10-174-579-570	Sequence 570, App	640	22	71.0	583	10	US-09-815-442-14002	Sequence 14002, A
568	22	71.0	425	9	US-10-174-582-570	Sequence 570, App	641	22	71.0	587	10	US-09-815-442-12082	Sequence 12082, A
569	22	71.0	425	9	US-10-174-588-570	Sequence 570, App	642	22	71.0	590	10	US-09-893-817-2	Sequence 2, Appl
570	22	71.0	425	9	US-10-175-739-570	Sequence 570, App	643	22	71.0	599	10	US-09-952-1133A-4	Sequence 4, Appl
571	22	71.0	425	9	US-10-175-740-570	Sequence 570, App	644	22	71.0	602	9	US-09-764-868-631	Sequence 631, App
572	22	71.0	425	9	US-10-175-743-570	Sequence 570, App	645	22	71.0	602	9	US-10-047-342-51	Sequence 51, Appl
573	22	71.0	425	9	US-10-176-488-570	Sequence 570, App	646	22	71.0	604	10	US-09-820-809-13	Sequence 13, Appl
574	22	71.0	425	9	US-10-176-492-570	Sequence 570, App	647	22	71.0	608	10	US-09-950-994-4	Sequence 4, Appl
575	22	71.0	425	9	US-10-176-747-570	Sequence 570, App	648	22	71.0	642	9	US-10-201-964-1	Sequence 1, Appl
576	22	71.0	425	9	US-10-176-750-570	Sequence 570, App	649	22	71.0	642	10	US-09-893-817-24	Sequence 24, Appl
577	22	71.0	425	9	US-10-176-985-570	Sequence 570, App	650	22	71.0	648	10	US-08-815-242-13995	Sequence 13995, A
578	22	71.0	425	9	US-10-176-987-570	Sequence 570, App	651	22	71.0	657	10	US-09-827-822-7	Sequence 7, Appl
579	22	71.0	425	9	US-10-176-991-570	Sequence 570, App	652	22	71.0	664	10	US-09-823-338A-48	Sequence 48, Appl
580	22	71.0	425	9	US-10-176-992-570	Sequence 570, App	653	22	71.0	669	10	US-09-801-196-28	Sequence 28, Appl
581	22	71.0	425	9	US-10-176-993-570	Sequence 570, App	654	22	71.0	690	9	US-10-068-059-10	Sequence 10, Appl
582	22	71.0	425	9	US-10-184-658-570	Sequence 570, App	655	22	71.0	709	9	US-10-068-059-8	Sequence 8, Appl
583	22	71.0	425	12	US-10-052-586-570	Sequence 570, App	656	22	71.0	709	10	US-09-815-242-5101	Sequence 5101, App
584	22	71.0	430	10	US-09-815-242-11315	Sequence 11315, A	657	22	71.0	720	9	US-10-047-412A-8	Sequence 8, Appl
585	22	71.0	432	10	US-09-815-242-10565	Sequence 10565, A	658	22	71.0	721	9	US-09-764-868B-731	Sequence 731, App
586	22	71.0	433	9	US-09-839-888-4	Sequence 4, Appl	659	22	71.0	724	9	US-10-068-059-12	Sequence 12, Appl
587	22	71.0	435	10	US-09-801-368-150	Sequence 150, App	660	22	71.0	731	9	US-09-738-626-4854	Sequence 4854, App
588	22	71.0	436	8	US-08-980-068B-19	Sequence 19, Appl	661	22	71.0	732	10	US-09-737-178-124	Sequence 124, App
589	22	71.0	437	9	US-09-943-123-14	Sequence 14, Appl	662	22	71.0	742	10	US-09-801-368B-26	Sequence 26, Appl
590	22	71.0	444	10	US-09-815-242-5482	Sequence 5482, App	663	22	71.0	746	9	US-10-068-059-6	Sequence 6, Appl
591	22	71.0	457	22	US-10-124-601-2	Sequence 2, Appl	664	22	71.0	746	9	US-09-982-107-4	Sequence 4, Appl
592	22	71.0	460	9	US-09-976-673-16	Sequence 16, Appl	665	22	71.0	756	10	US-09-771-61A-101	Sequence 101, App
593	22	71.0	470	9	US-09-976-673-18	Sequence 18, Appl	666	22	71.0	764	9	US-09-981-353-59	Sequence 59, Appl
594	22	71.0	471	9	US-09-738-626-3529	Sequence 3529, App	667	22	71.0	764	10	US-09-989-919-124	Sequence 124, App
595	22	71.0	473	10	US-09-815-242-12584	Sequence 12584, A	668	22	71.0	764	10	US-09-818-447-1	Sequence 1, Appl
596	22	71.0	478	9	US-09-859-888-2	Sequence 2, Appl	669	22	71.0	781	10	US-09-815-242-11269	Sequence 425, App
597	22	71.0	501	10	US-09-815-242-11300	Sequence 11300, A	670	22	71.0	819	10	US-09-833-290-425	Sequence 1819, App
598	22	71.0	502	10	US-09-323-998B-56	Sequence 56, Appl	671	22	71.0	831	9	US-09-738-487-1819	Sequence 1819, App
599	22	71.0	508	10	US-09-721-161A-102	Sequence 102, App	672	22	71.0	831	9	US-09-902-941-1819	Sequence 1819, App
600	22	71.0	511	10	US-09-917-265-44	Sequence 44, Appl	673	22	71.0	831	9	US-09-845-626-1819	Sequence 27, Appl
601	22	71.0	511	10	US-09-917-265-67	Sequence 67, Appl	674	22	71.0	854	10	US-10-128-870-27	Sequence 11, Appl
602	22	71.0	515	9	US-09-738-626-4135	Sequence 4135, App	675	22	71.0	863	10	US-09-946-339-11	Sequence 5, Appl
603	22	71.0	517	10	US-09-796-138-21	Sequence 21, Appl	676	22	71.0	872	10	US-09-813-148-5	Sequence 5, Appl

677	22	71.0	885	10	US-09-815-242-5090	Sequence 5090, App	750	22	71.0	1177	10	US-09-873-873-28	Sequence 28, Appl
678	22	71.0	887	10	US-09-771-161A-192	Sequence 192, App	751	22	71.0	1177	10	US-09-873-873-34	Sequence 34, Appl
679	22	71.0	887	10	US-09-771-161A-193	Sequence 193, App	752	22	71.0	1178	10	US-09-851-194-2	Sequence 2, Appl
680	22	71.0	896	9	US-09-903-170C-5	Sequence 5, Appl	753	22	71.0	1193	10	US-09-873-873-30	Sequence 30, Appl
681	22	71.0	896	10	US-09-903-180B-5	Sequence 5, Appl	754	22	71.0	1221	10	US-09-738-626-513	Sequence 513, Ap
682	22	71.0	896	10	US-09-903-171A-5	Sequence 5, Appl	755	22	71.0	1221	10	US-09-919-891-2	Sequence 2, Appl
683	22	71.0	896	10	US-09-903-188A-5	Sequence 5, Appl	756	22	71.0	1229	10	US-09-815-244-5818	Sequence 5818, Ap
684	22	71.0	896	10	US-09-903-123A-5	Sequence 5, Appl	757	22	71.0	1229	10	US-09-815-244-1946	Sequence 12946, A
685	22	71.0	896	10	US-09-903-325A-5	Sequence 5, Appl	758	22	71.0	1241	9	US-09-759-622-8	Sequence 2, Appl
686	22	71.0	908	10	US-09-815-242-11847	Sequence 11847, A	759	22	71.0	1241	9	US-10-047-542-89	Sequence 89, Appl
687	22	71.0	912	10	US-09-817-664-2	Sequence 2, Appl	760	22	71.0	1274	9	US-10-020-215-2	Sequence 2, Appl
688	22	71.0	922	10	US-09-738-626-5476	Sequence 5476, App	761	22	71.0	1356	10	US-09-801-366-306	Sequence 306, App
689	22	71.0	922	10	US-09-886-468-19	Sequence 19, App	762	22	71.0	1352	9	US-09-738-626-4321	Sequence 4321, Ap
690	22	71.0	924	9	US-10-108-605-137	Sequence 137, App	763	22	71.0	1607	10	US-09-938-275-10	Sequence 10, Appl
691	22	71.0	941	9	US-09-993-598-353	Sequence 353, App	764	22	71.0	1609	10	US-09-938-275-11	Sequence 11, Appl
692	22	71.0	941	9	US-09-989-293A-353	Sequence 353, App	765	22	71.0	1633	9	US-09-738-626-6894	Sequence 6894, Ap
693	22	71.0	941	9	US-10-042-141-60	Sequence 60, Appl	766	22	71.0	1809	10	US-09-822-633-2	Sequence 2, Appl
694	22	71.0	941	9	US-09-989-735-353	Sequence 353, App	767	22	71.0	2249	9	US-09-865-557-2	Sequence 4, Appl
695	22	71.0	941	9	US-09-990-444-353	Sequence 353, App	768	22	71.0	26936	9	US-09-759-5088-2	Sequence 2, Appl
696	22	71.0	941	9	US-09-989-730-353	Sequence 353, App	769	22	67.7	10	US-09-729-873-117	Sequence 117, App	
697	22	71.0	941	9	US-09-990-436-353	Sequence 353, App	770	21	67.7	13	9	US-09-774-639-277	Sequence 277, App
698	22	71.0	941	9	US-09-991-181-353	Sequence 353, App	771	21	67.7	16	9	US-09-840-277-86	Sequence 82, Appl
699	22	71.0	941	9	US-09-993-687-353	Sequence 353, App	772	21	67.7	16	9	US-09-840-277-86	Sequence 86, Appl
700	22	71.0	941	9	US-09-989-734-353	Sequence 353, App	773	21	67.7	21	10	US-09-853-830-108	Sequence 108, App
701	22	71.0	941	9	US-10-028-072-464	Sequence 464, App	774	21	67.7	32	10	US-09-861-624-28	Sequence 25, Appl
702	22	71.0	941	9	US-09-993-653-353	Sequence 353, App	775	21	67.7	36	10	US-09-864-761-3869	Sequence 3869, A
703	22	71.0	941	9	US-09-993-667-353	Sequence 353, App	776	21	67.7	36	10	US-09-733-524-9	Sequence 9, Appl
704	22	71.0	941	9	US-10-121-049-464	Sequence 464, App	777	21	67.7	46	10	US-09-864-761-47826	Sequence 47826, A
705	22	71.0	941	9	US-10-123-904-464	Sequence 464, App	778	21	67.7	47	10	US-09-864-761-49041	Sequence 49041, A
706	22	71.0	941	9	US-10-140-470-464	Sequence 464, App	779	21	67.7	47	12	US-10-001-8179-114	Sequence 114, App
707	22	71.0	941	9	US-09-990-438-353	Sequence 353, App	780	21	67.7	53	9	US-09-843-676-20	Sequence 20, Appl
708	22	71.0	941	9	US-09-990-562-353	Sequence 353, App	781	21	67.7	53	9	US-09-765-253-20	Sequence 20, Appl
709	22	71.0	941	9	US-09-997-428-353	Sequence 353, App	782	21	67.7	53	9	US-09-438-486-20	Sequence 20, Appl
710	22	71.0	941	9	US-09-997-666-353	Sequence 353, App	783	21	67.7	58	10	US-09-864-761-40567	Sequence 40567, A
711	22	71.0	941	9	US-10-175-746-464	Sequence 464, App	784	21	67.7	61	8	US-08-808-031A-19	Sequence 19, Appl
712	22	71.0	941	9	US-10-176-918-464	Sequence 464, App	785	21	67.7	63	10	US-09-864-761-37193	Sequence 37193, A
713	22	71.0	941	10	US-10-176-921-464	Sequence 464, App	786	21	67.7	68	9	US-09-895-913A-272	Sequence 272, App
714	22	71.0	941	10	US-09-726-643-60	Sequence 60, Appl	787	21	67.7	70	10	US-09-974-879-191	Sequence 191, App
715	22	71.0	941	10	US-09-989-723-353	Sequence 353, App	788	21	67.7	70	10	US-09-864-761-34690	Sequence 34690, A
716	22	71.0	941	10	US-09-989-723-353	Sequence 353, App	789	21	67.7	71	10	US-09-864-761-36979	Sequence 36979, A
717	22	71.0	941	10	US-09-989-279-353	Sequence 353, App	790	21	67.7	72	10	US-09-925-301-1505	Sequence 1505, Ap
718	22	71.0	941	10	US-09-989-727-353	Sequence 353, App	791	21	67.7	72	10	US-09-808-212A-8	Sequence 8, Appl
719	22	71.0	941	10	US-09-989-731-353	Sequence 353, App	792	21	67.7	72	10	US-09-808-212A-10	Sequence 10, Appl
720	22	71.0	941	10	US-09-989-732-353	Sequence 353, App	793	21	67.7	73	10	US-09-939-980-416	Sequence 416, App
721	22	71.0	941	10	US-09-991-073-353	Sequence 353, App	794	21	67.7	75	10	US-09-808-212A-18	Sequence 18, Appl
722	22	71.0	941	10	US-09-990-442-353	Sequence 353, App	795	21	67.7	77	10	US-09-939-980-511	Sequence 511, App
723	22	71.0	941	10	US-09-991-163-353	Sequence 353, App	796	21	67.7	82	10	US-09-808-212A-2	Sequence 2, Appl
724	22	71.0	941	10	US-09-993-604-353	Sequence 353, App	797	21	67.7	83	9	US-10-042-141-84	Sequence 84, Appl
725	22	71.0	941	10	US-09-990-456-353	Sequence 353, App	798	21	67.7	83	10	US-09-726-643-84	Sequence 84, Appl
726	22	71.0	941	10	US-09-989-721-353	Sequence 353, App	799	21	67.7	84	9	US-09-974-879-289	Sequence 289, Appl
727	22	71.0	961	10	US-09-220-091-3	Sequence 3, Appl	800	21	67.7	85	10	US-09-730-611-35	Sequence 35, Appl
728	22	71.0	978	9	US-09-938-901-8	Sequence 8, Appl	801	21	67.7	87	10	US-09-867-550-36	Sequence 36, Appl
729	22	71.0	979	10	US-09-903-187A-5	Sequence 5, Appl	802	21	67.7	91	9	US-09-764-868-691	Sequence 691, App
730	22	71.0	990	12	US-10-047-676A-7	Sequence 7, Appl	803	21	67.7	95	9	US-09-764-868-784	Sequence 784, App
731	22	71.0	993	10	US-09-815-242-5809	Sequence 5809, App	804	21	67.7	95	9	US-09-764-868-1188	Sequence 1188, Ap
732	22	71.0	993	10	US-09-881-752A-80	Sequence 80, Appl	805	21	67.7	105	9	US-09-738-626-3621	Sequence 3621, Ap
733	22	71.0	999	9	US-09-895-913A-226	Sequence 226, App	806	21	67.7	109	10	US-09-925-299-1043	Sequence 1043, Ap
734	22	71.0	1002	10	US-09-815-242-12899	Sequence 12899, A	807	21	67.7	109	10	US-09-220-206A-7	Sequence 7, Appl
735	22	71.0	1002	10	US-09-815-242-13158	Sequence 13158, A	808	21	67.7	115	9	US-09-738-626-40474	Sequence 40474, A
736	22	71.0	1040	10	US-09-765-772-118	Sequence 118, App	809	21	67.7	115	10	US-09-265-995-352	Sequence 352, App
737	22	71.0	1042	10	US-09-815-242-13345	Sequence 13345, A	810	21	67.7	117	10	US-09-864-761-37175	Sequence 37175, A
738	22	71.0	1042	10	US-09-525-697-2	Sequence 2, Appl	811	21	67.7	119	8	US-08-945-456A-1	Sequence 1, Appl
739	22	71.0	1155	10	US-09-756-643-2	Sequence 2, Appl	812	21	67.7	119	9	US-09-880-708-13	Sequence 13, Appl
740	22	71.0	1156	10	US-09-826-660-15	Sequence 15, Appl	813	21	67.7	119	9	US-09-068-253-2	Sequence 2, Appl
741	22	71.0	1160	10	US-09-815-242-5480	Sequence 5480, Ap	814	21	67.7	119	10	US-09-864-761-43653	Sequence 43653, A
742	22	71.0	1168	10	US-09-815-242-12151	Sequence 12151, A	815	21	67.7	120	9	US-09-945-182-4	Sequence 4, Appl
743	22	71.0	1169	10	US-09-815-242-13448	Sequence 13448, A	816	21	67.7	122	12	US-10-115-406-7	Sequence 7, Appl
744	22	71.0	1169	10	US-09-945-265-2	Sequence 2, Appl	817	21	67.7	123	9	US-09-859-211-36	Sequence 36, Appl
745	22	71.0	1170	10	US-09-815-242-13693	Sequence 13693, A	818	21	67.7	123	10	US-09-815-459-7	Sequence 7, Appl
746	22	71.0	1177	10	US-09-873-873-12	Sequence 12, Appl	819	21	67.7	131	9	US-09-738-626-5916	Sequence 5916, Ap
747	22	71.0	1177	10	US-09-873-873-10	Sequence 10, Appl	820	21	67.7	135	10	US-09-764-868-1098	Sequence 1098, Ap
748	22	71.0	1177	10	US-09-873-873-14	Sequence 14, Appl	821	21	67.7	140	10	US-09-730-617-46	Sequence 46, Appl
749	22	71.0	1177	10	US-09-873-873-26	Sequence 26, Appl	822	21	67.7	142	10	US-09-797-481-2	Sequence 2, Appl

823	21	67.7	142	10	US-09-844-736-4	Sequence 4, Appl1	896	21	67.7	312	10	US-09-815-242-5267	Sequence 5667, Ap
824	21	67.7	144	9	US-09-929-230-8	Sequence 8, Appl1	897	21	67.7	316	10	US-09-462-846-5	Sequence 5, Appl1
825	21	67.7	144	10	US-09-730-617-45	Sequence 45, Appl	898	21	67.7	321	10	US-09-945-182-26	Sequence 26, Appl
826	21	67.7	154	9	US-10-090-185-11	Sequence 24, Appl	899	21	67.7	321	10	US-09-886-055-259	Sequence 259, Ap
827	21	67.7	155	10	US-09-425-021-24	Sequence 11, Appl	900	21	67.7	323	10	US-09-886-055-77	Sequence 77, Ap
828	21	67.7	174	9	US-09-764-868-1131	Sequence 613, Ap	901	21	67.7	323	10	US-09-771-730-8	Sequence 8, Appl1
829	21	67.7	175	10	US-09-925-297-613	Sequence 293, Ap	902	21	67.7	323	10	US-09-771-730-16	Sequence 16, Appl
830	21	67.7	177	10	US-09-741-669-293	Sequence 317, Ap	903	21	67.7	323	10	US-09-771-730-18	Sequence 18, Appl
831	21	67.7	177	10	US-09-912-020-317	Sequence 5164, Ap	904	21	67.7	323	10	US-09-771-730-18	Sequence 31, Appl
832	21	67.7	177	10	US-09-815-242-5164	Sequence 10344, A	905	21	67.7	328	9	US-09-735-055-31	Sequence 13188, A
833	21	67.7	177	10	US-09-815-242-11112	Sequence 11112, A	906	21	67.7	342	10	US-09-849-9808-1	Sequence 1, Appl1
834	21	67.7	177	10	US-09-815-242-11112	Sequence 14103, A	907	21	67.7	343	10	US-09-864-761-16084	Sequence 36084, A
835	21	67.7	177	10	US-09-815-242-14103	Sequence 11, Appl	908	21	67.7	346	10	US-09-791-961-5	Sequence 5, Appl1
836	21	67.7	179	9	US-10-223-047-11	Sequence 186, Ap	909	21	67.7	350	10	US-09-791-961-5	Sequence 1, Appl1
837	21	67.7	180	10	US-09-765-272-186	Sequence 12316, A	910	21	67.7	352	10	US-09-815-242-11287	Sequence 11287, A
838	21	67.7	182	10	US-09-815-242-12316	Sequence 10, Appl	911	21	67.7	353	10	US-09-815-242-5323	Sequence 5323, Ap
839	21	67.7	182	10	US-09-975-374A-10	Sequence 5833, Ap	912	21	67.7	354	9	US-10-121-851-38	Sequence 38, Appl
840	21	67.7	183	10	US-09-815-242-5833	Sequence 2, Appl1	913	21	67.7	357	10	US-09-740-046-8	Sequence 8, Appl1
841	21	67.7	190	10	US-09-877-964-2	Sequence 4, Appl1	914	21	67.7	358	10	US-09-815-242-12380	Sequence 12380, A
842	21	67.7	190	10	US-09-877-964-4	Sequence 2, Appl1	915	21	67.7	358	10	US-09-925-637-2	Sequence 2, Appl1
843	21	67.7	190	10	US-09-905-558C-2	Sequence 32, Appl	916	21	67.7	359	12	US-10-072-152-2	Sequence 4, Appl1
844	21	67.7	197	9	US-10-140-293-32	Sequence 1216, Ap	917	21	67.7	360	9	US-09-738-626-4558	Sequence 4958, Ap
845	21	67.7	198	9	US-09-796-692-1216	Sequence 1679, Ap	918	21	67.7	360	10	US-09-815-242-11186	Sequence 11186, A
846	21	67.7	199	9	US-09-796-692-1675	Sequence 2233, Ap	919	21	67.7	361	9	US-09-738-626-4650	Sequence 4650, Ap
847	21	67.7	199	9	US-09-796-692-2234	Sequence 31, Appl	920	21	67.7	363	10	US-09-815-242-12579	Sequence 12579, A
848	21	67.7	199	9	US-10-140-293-31	Sequence 13428, A	921	21	67.7	365	10	US-09-815-242-11917	Sequence 11917, A
849	21	67.7	204	10	US-09-815-242-13428	Sequence 13428, A	922	21	67.7	371	10	US-09-925-300-1681	Sequence 1681, Ap
850	21	67.7	204	10	US-09-815-242-13687	Sequence 13687, A	923	21	67.7	373	10	US-09-815-398-36	Sequence 36, Appl
851	21	67.7	206	10	US-09-935-428A-10	Sequence 10, Appl	924	21	67.7	373	10	US-09-905-173-26	Sequence 26, Appl
852	21	67.7	210	10	US-09-997-165-6	Sequence 6, Appl	925	21	67.7	380	10	US-09-815-242-12770	Sequence 12770, A
853	21	67.7	220	9	US-09-860-670-129	Sequence 129, Ap	926	21	67.7	381	9	US-09-738-626-6596	Sequence 6596, Ap
854	21	67.7	220	9	US-09-764-868-949	Sequence 949, Ap	927	21	67.7	382	10	US-09-815-242-5342	Sequence 5342, Ap
855	21	67.7	220	9	US-09-764-868-1331	Sequence 1231, Ap	928	21	67.7	382	10	US-09-815-242-13912	Sequence 13912, A
856	21	67.7	220	9	US-09-770-940-5	Sequence 5, Appl1	929	21	67.7	384	9	US-09-924-400-326	Sequence 324, Ap
857	21	67.7	220	10	US-09-765-272-78	Sequence 78, Appl	930	21	67.7	384	9	US-09-924-400-326	Sequence 324, Ap
858	21	67.7	221	9	US-10-170-102-9	Sequence 9, Appl	931	21	67.7	384	9	US-09-924-400-332	Sequence 332, Ap
859	21	67.7	224	10	US-09-765-213A-11	Sequence 11, Appl	932	21	67.7	384	10	US-09-924-400-334	Sequence 334, Ap
860	21	67.7	240	10	US-09-945-182-30	Sequence 30, Appl	933	21	67.7	384	10	US-09-825-301-8	Sequence 8, Appl1
861	21	67.7	240	10	US-09-997-165-2	Sequence 2, Appl1	934	21	67.7	384	10	US-09-815-242-5184	Sequence 5184, Ap
862	21	67.7	242	10	US-09-815-242-11694	Sequence 11694, A	935	21	67.7	384	10	US-09-815-242-11196	Sequence 11196, A
863	21	67.7	248	9	US-09-738-626-3952	Sequence 3952, Ap	936	21	67.7	384	10	US-09-810-936-304	Sequence 304, Ap
864	21	67.7	248	10	US-09-725-178-2	Sequence 2, Appl1	937	21	67.7	384	10	US-09-810-936-326	Sequence 326, Ap
865	21	67.7	258	10	US-09-839-185-10	Sequence 10, Appl	938	21	67.7	384	10	US-09-810-936-332	Sequence 332, Ap
866	21	67.7	258	9	US-09-870-759-14	Sequence 14, Appl	939	21	67.7	384	10	US-09-810-936-334	Sequence 334, Ap
867	21	67.7	261	10	US-09-815-242-5376	Sequence 5376, Ap	940	21	67.7	384	10	US-09-429-573-304	Sequence 304, Ap
868	21	67.7	263	10	US-09-945-182-32	Sequence 32, Appl	941	21	67.7	388	9	US-09-945-182-24	Sequence 24, Appl
869	21	67.7	263	10	US-09-859-882-4	Sequence 4, Appl1	942	21	67.7	390	10	US-09-815-242-12402	Sequence 12402, A
870	21	67.7	264	10	US-09-726-899-1	Sequence 1, Appl1	943	21	67.7	392	9	US-09-738-626-4635	Sequence 4635, Ap
871	21	67.7	266	10	US-09-726-899-9	Sequence 9, Appl1	944	21	67.7	394	9	US-09-924-400-336	Sequence 336, Ap
872	21	67.7	266	10	US-09-815-242-10197	Sequence 10197, A	945	21	67.7	394	10	US-09-810-936-332	Sequence 332, Ap
873	21	67.7	266	10	US-09-815-242-12291	Sequence 12291, A	946	21	67.7	411	9	US-09-815-242-12007	Sequence 12007, A
874	21	67.7	268	10	US-09-925-637-60	Sequence 60, Appl	947	21	67.7	411	9	US-09-945-182-28	Sequence 28, Appl
875	21	67.7	276	9	US-09-764-868-811	Sequence 811, Ap	948	21	67.7	419	10	US-09-815-242-13330	Sequence 13330, A
876	21	67.7	276	9	US-09-738-626-5519	Sequence 5519, Ap	949	21	67.7	422	10	US-09-905-119-2	Sequence 2, Appl1
877	21	67.7	282	10	US-09-815-242-14057	Sequence 14057, A	950	21	67.7	423	10	US-09-798-029-8	Sequence 8, Appl1
878	21	67.7	282	12	US-10-116-253-7	Sequence 7, Appl1	951	21	67.7	425	10	US-09-828-313-29	Sequence 29, Appl
879	21	67.7	289	10	US-09-815-242-11940	Sequence 11940, A	952	21	67.7	425	10	US-09-828-313-29	Sequence 29, Appl
880	21	67.7	290	10	US-09-815-242-5743	Sequence 5743, Ap	953	21	67.7	434	1	US-08-325-278-3	Sequence 3, Appl1
881	21	67.7	293	21	US-09-815-242-12742	Sequence 12742, A	954	21	67.7	434	9	US-09-738-626-6610	Sequence 6610, Ap
882	21	67.7	294	9	US-09-945-182-2	Sequence 2, Appl1	955	21	67.7	435	10	US-09-731-8266-4	Sequence 4, Appl1
883	21	67.7	296	10	US-09-960-472-1	Sequence 1, Appl1	956	21	67.7	441	10	US-09-815-242-5702	Sequence 5702, Ap
884	21	67.7	299	10	US-09-771-730-85	Sequence 85, Appl	957	21	67.7	441	10	US-09-925-301-1303	Sequence 1303, Ap
885	21	67.7	299	10	US-09-771-730-90	Sequence 90, Appl	958	21	67.7	443	10	US-09-929-060-1	Sequence 1, Appl1
886	21	67.7	302	10	US-09-771-730-94	Sequence 94, Appl	959	21	67.7	443	10	US-09-929-060-2	Sequence 2, Appl1
887	21	67.7	302	10	US-09-854-122-19	Sequence 19, Appl	960	21	67.7	446	10	US-09-747-116-2	Sequence 12, Appl
888	21	67.7	302	10	US-09-771-730-61	Sequence 61, Appl	961	21	67.7	447	9	US-10-081-309-2	Sequence 9, Appl1
889	21	67.7	303	10	US-09-771-730-92	Sequence 92, Appl	962	21	67.7	451	9	US-09-910-186A-12	Sequence 12, Appl
890	21	67.7	305	1	US-08-325-278-1	Sequence 1, Appl1	963	21	67.7	452	10	US-09-764-870-339	Sequence 339, Ap
891	21	67.7	305	10	US-09-933-92A-2	Sequence 2, Appl1	964	21	67.7	452	10	US-09-866-356-4	Sequence 4, Appl1
892	21	67.7	309	10	US-09-801-568-286	Sequence 286, Appl	965	21	67.7	454	9	US-09-738-626-5211	Sequence 5211, Ap
893	21	67.7	310	9	US-09-738-626-3669	Sequence 3669, Ap	966	21	67.7	459	10	US-09-866-987-8	Sequence 8, Appl1
894	21	67.7	312	9	US-09-738-626-6812	Sequence 6812, Ap	967	21	67.7	460	10	US-09-841-880-2	Sequence 35, Appl
895	21	67.7	312	10	US-09-788-626-21	Sequence 21, Appl	968	21	67.7	461	10	US-09-862-027-35	Sequence 35, Appl
										463	9	US-09-738-626-3970	Sequence 3970, Ap

```

569 21 67.7 467 10 US-09-929-060-3 Sequence 3, Appli
570 21 67.7 467 12 US-10-083-452-11 Sequence 11, Appl
571 21 67.7 477 9 US-09-738-626-6985 Sequence 6985, Ap
572 21 67.7 479 10 US-09-881-752A-14 Sequence 14, Appl
573 21 67.7 481 10 US-09-817-464-8 Sequence 8, Appl
574 21 67.7 485 10 US-09-815-242-5503 Sequence 5503, Ap
575 21 67.7 491 9 US-09-885-913A-106 Sequence 106, App
576 21 67.7 491 10 US-09-815-242-12492 Sequence 12492, A
577 21 67.7 491 10 US-09-815-242-12864 Sequence 12864, A
578 21 67.7 494 10 US-09-815-242-11055 Sequence 11055, A
579 21 67.7 495 9 US-09-880-708-10 Sequence 10, Appl
580 21 67.7 496 9 US-09-784-249-2 Sequence 2, Appl
581 21 67.7 501 8 US-08-981-490B-1 Sequence 1, Appl
582 21 67.7 501 10 US-09-730-772-13 Sequence 13, Appl
583 21 67.7 501 10 US-09-735-849-13 Sequence 13, Appl
584 21 67.7 502 9 US-09-813-398-37 Sequence 37, Appl
585 21 67.7 502 10 US-09-815-242-11205 Sequence 11205, A
586 21 67.7 504 10 US-09-815-242-12537 Sequence 12537, A
587 21 67.7 509 9 US-09-258-031B-75 Sequence 75, Appl
588 21 67.7 511 9 US-10-121-032-21 Sequence 21, Appl
589 21 67.7 511 9 US-09-941-947A-22 Sequence 22, Appl
590 21 67.7 511 10 US-09-934-903-16 Sequence 16, Appl
591 21 67.7 511 10 US-09-934-868-76 Sequence 76, Appl
592 21 67.7 526 9 US-09-738-626-6454 Sequence 6454, Ap
593 21 67.7 528 9 US-10-121-498-12 Sequence 12, Appl
594 21 67.7 529 9 US-09-924-400-324 Sequence 324, App
595 21 67.7 529 9 US-08-258-031B-16 Sequence 16, Appl
596 21 67.7 529 9 US-08-258-031B-20 Sequence 20, Appl
597 21 67.7 529 10 US-09-810-936-324 Sequence 324, App
598 21 67.7 533 9 US-09-905-291A-332 Sequence 332, App
599 21 67.7 533 9 US-09-902-853-332 Sequence 332, App
1000 21 67.7 533 9 US-09-907-824-332 Sequence 332, App

```

ALIGNMENTS

```

RESULT 1
US-09-816-825-9
; Sequence 9, Application US/09816825
; Patient No. US20010051370A1
; GENERAL INFORMATION:
; APPLICANT: Bistup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CON
; CURRENT APPLICATION NUMBER: US/09/816,825
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/045,284
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-825-9

Query Match          100.0%; Score 31; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-9

```

```

Query Match          100.0%; Score 31; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VRYEDL 6
DB 127 VRYEDL 132

```

```

RESULT 3
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patient No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8

```

```

Query Match          100.0%; Score 31; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VRYEDL 6
DB 129 VRYEDL 134

```

```

RESULT 4
US-09-927-602-10
; Sequence 10, Application US/09927602
; Patient No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0

```

SEQ ID NO 10
LENGTH: 179
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-10

Query Match 100.0%; Score 31; DB 10; Length 179;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 137 VRYEDL 142

RESULT 5
US-10-007-262-1
Sequence 1, Application US/10007262
Patent No. US20020164748A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107C1P
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 31; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 274 VRYEDL 279

RESULT 6
US-09-816-825-2
Sequence 2, Application US/09816825
Patent No. US20010051370A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-825-2

Query Match 100.0%; Score 31; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 274 VRYEDL 279

RESULT 7
US-09-927-602-5
Sequence 5, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
Dystrophy
FILE REFERENCE: P-LT 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-927-602-5

Query Match 100.0%; Score 31; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 294 VRYEDL 299

RESULT 8
US-10-212-933-2
Sequence 2, Application US/10212933
Publication No. US20030008366A1
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadamatsu, Kenji
APPLICANT: Kannagi, Reiji
APPLICANT: Habuchi, Osami
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGALACTOSAMINE-6-O-
SULFOTRANSFERASE AND
TITLE OF INVENTION: DNA ENCODING THE SAME
FILE REFERENCE: TOYAMA1,001AUS
CURRENT APPLICATION NUMBER: US/10/212,933
CURRENT FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: US/09/263,023
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: JP 10-54007
PRIOR FILING DATE: 1998-03-05
PRIOR APPLICATION NUMBER: JP 10-177844
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-10-212-933-2

Query Match 100.0%; Score 31; DB 9; Length 483;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 376 VREDEL 381

```
RESULT 9
US-10-212-933-4
; Sequence 4, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Muramatsu, Kenji
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA41.001AUS
; CURRENT APPLICATION NUMBER: US/10/212.933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263.023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-933-4
```

```
Query Match 100.0%; Score 31; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VREDEL 6
Db 377 VREDEL 382
```

```
RESULT 10
US-09-833-790-255
; Sequence 255, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodge, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Radoch
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833.790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-255
```

```
Query Match 100.0%; Score 31; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VREDEL 6
Db 377 VREDEL 382
```

Db 424 VREDEL 429

```
RESULT 11
US-09-927-602-11
; Sequence 11, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LD 4852
; CURRENT APPLICATION NUMBER: US/09/927.602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638.211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-11
```

```
Query Match 90.3%; Score 28; DB 10; Length 174;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VREDEL 6
Db 132 VREDEL 137
```

```
RESULT 12
US-10-000-864-6
; Sequence 6, Application US/10000864
; Patent No. US20020146798A1
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK1 PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CRI-085CPC
; CURRENT APPLICATION NUMBER: US/10/000.864
; CURRENT FILING DATE: 2001-10-31
; EARLIER APPLICATION NUMBER: 09/423.890
; EARLIER FILING DATE: 2000-06-03
; EARLIER APPLICATION NUMBER: PCT/US99/05556
; EARLIER FILING DATE: 1999-03-15
; EARLIER APPLICATION NUMBER: USSN 60/078.153
; EARLIER FILING DATE: 1998-03-16
; EARLIER APPLICATION NUMBER: USSN 60/099.165
; EARLIER FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-864-6
```

```
Query Match 90.3%; Score 28; DB 12; Length 626;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VREDEL 6
Db 64 VREDEL 69
```

```
RESULT 13
US-10-000-864-12
; Sequence 12, Application US/10000864
```

Patent No. US20020146798A1
GENERAL INFORMATION:
APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK1 PROTEIN AND NUCLEIC ACID MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/10/000,864
EARLIER FILING DATE: 2001-10-31
EARLIER APPLICATION NUMBER: 09/423,890
EARLIER FILING DATE: 2000-06-03
EARLIER APPLICATION NUMBER: PCT/US99/05556
EARLIER FILING DATE: 1999-03-15
EARLIER APPLICATION NUMBER: USSN 60/078,153
EARLIER FILING DATE: 1998-03-16
EARLIER APPLICATION NUMBER: USSN 60/099,165
EARLIER FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-10-000-864-12

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 626;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VYEDL 6
DB 64 VYEDV 69

RESULT 14
US-09-864-761-35570
Sequence 35570, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 35570
LENGTH: 69
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096700.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.9
OTHER INFORMATION: EST_HUMAN HIT: D44816.1, EVALU 5.00e-35
OTHER INFORMATION: SWISSPROT HIT: P40421, EVALU 7.00e-04
US-09-864-761-35570

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 56 RYEDL 60

RESULT 15
US-09-864-761-49114
Sequence 49114, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 49114
LENGTH: 69
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096700.14
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
OTHER INFORMATION: SWISSPROT HIT: P40421, EVALU 7.00e-04
OTHER INFORMATION: EST_HUMAN HIT: D44816.1, EVALU 8.00e-35
US-09-864-761-49114

Query Match 87.1%; Score 27; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
|||
Db 56 RYEDL 60

RESULT 16
US-09-738-626-3600
Sequence 3600, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIRO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3600
LENGTH: 157
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3600

Query Match 87.1%; Score 27; DB 9; Length 157;
Best Local Similarity 83.3%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||
Db 25 VRYEDL 30

RESULT 17
US-09-927-602-6
Sequence 6, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Homo sapien
US-09-927-602-6

Query Match 87.1%; Score 27; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||
Db 127 VRYEDL 132

RESULT 18
US-09-927-602-7
Sequence 7, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-7

Query Match 87.1%; Score 27; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||
Db 127 VRYEDL 132

RESULT 19
US-09-361-741-2


```
; Sequence 2, Application US/09361741
; Patient No. US20020048784A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, PETER L
; APPLICANT: ROSEN, CRAIG A
; APPLICANT: HE, WEI WU
; TITLE OF INVENTION: PROSTATIC GROWTH FACTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,741
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,607
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-329
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)994-1700
; TELEFAX: (201)994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-361-741-2

Query Match      87.1%; Score 27; DB 10; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      46 RYEDL 50

RESULT 20
; Sequence 188, Application US/09981353
; Patient No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W,
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 188
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 522433CD1
; US-09-981-353-188
```

```
Query Match      87.1%; Score 27; DB 9; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      59 RYEDL 63

RESULT 21
; Sequence 6, Application US/09276600
; Patient No. US20010010908A1
; GENERAL INFORMATION:
; APPLICANT: Patricia Billing-Medel
; APPLICANT: Maurice Cohen
; APPLICANT: Tracey L. Colpitts
; APPLICANT: Julian Gordon
; APPLICANT: Edward N. Granados
; APPLICANT: John C. Russell
; APPLICANT: Stephen D. Stroupe
; TITLE OF INVENTION: Reagents and Methods useful for
; TITLE OF INVENTION: Detecting Disease of the Prostate
; FILE REFERENCE: 6397.US.01
; CURRENT APPLICATION NUMBER: US/09/276,600
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-276-600-6

Query Match      87.1%; Score 27; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      59 RYEDL 63

RESULT 22
; Sequence 14, Application US/09966546
; Patient No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-966-546-14

Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RYED 5
        |||||
```

Db 262 VYED 266

RESULT 23

```
US-09-966-545-14
; Sequence 14, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20020172999A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-343-14
```

```
Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYED 5

Db 262 VYED 266

RESULT 24

```
US-09-965-212-14
; Sequence 14, Application US/09965212
; Publication No. US20030003462A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Vernet, Corine
; TITLE OF INVENTION: No. US20030003462A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-14
```

```
Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYED 5

Db 262 VYED 266

RESULT 25

```
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
```

GENERAL INFORMATION:

```
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-4
```

```
Query Match      87.1%; Score 27; DB 10; Length 390;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYEDL 6

Db 272 VYEDL 277

RESULT 26

```
US-09-927-602-2
; Sequence 2, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-2
```

```
Query Match      87.1%; Score 27; DB 10; Length 395;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYEDL 6

Db 271 VYEDL 276

RESULT 27

```
US-09-927-602-3
; Sequence 3, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: VARIANT
; LOCATION: (1)...(395)
; OTHER INFORMATION: Xaa = any amino acid
US-09-927-602-3

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 395;
Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYEDL 6
DB 271 RYEDL 276

RESULT 28
US-09-815-242-12541
; Sequence 12541, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12541
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12541

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 396;
Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 29
US-09-815-242-12900
; Sequence 12900, Application US/09815242
```

```

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12900

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 396;
Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 30
US-09-741-669-300
; Sequence 300, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 300
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-300

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 439;
Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 RYEDL 6
|||||
DB 112 RYEDL 116

RESULT 31

US-09-815-242-10284
; Sequence 10284, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zykand, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: EXTRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10284
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10284

Query Match 87.1%; Score 27; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
|||||
DB 112 RYEDL 116

US-09-907-509-2
; Sequence 2, Application US/09907509
; Patent No. US20020090705A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 62088, A NOVEL HUMAN NUCLEOSIDE
; FILE REFERENCE: MNT-177
; CURRENT APPLICATION NUMBER: US/09/907,509
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/216385
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 604
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-907-509-2

Query Match 87.1%; Score 27; DB 10; Length 604;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
|||||
DB 323 RYEDL 327

RESULT 33

US-09-771-467C-4
; Sequence 4, Application US/09771467C
; Patent No. US20020177549A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Yuling
; APPLICANT: Xiomel, Xu
; TITLE OF INVENTION: Semaphorin X1 Polypeptides
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/771,467C
; FILING DATE: 26-Jan-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: EXEL98-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-771-467C-4

Query Match 87.1%; Score 27; DB 9; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYED 5
|||||
DB 415 RYED 419

US-09-808-602-59

; Sequence 59, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Rumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S

```

; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIORITY FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-59
```

```

Query Match      87.1%; Score 27; DB 9; Length 693;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VRYED 5
        |||||
DB      254 VRYED 258
```

```

RESULT 35
US-09-808-602-60
; Sequence 60, Application US/09808602
; Patent No. US20020155115A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIORITY FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-60
```

```

Query Match      87.1%; Score 27; DB 9; Length 773;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VRYED 5
        |||||
DB      329 VRYED 333
```

```

RESULT 36
US-09-808-602-4
; Sequence 4, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
```

```

; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIORITY FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 815
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-4
```

```

Query Match      87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VRYED 5
        |||||
DB      376 VRYED 380
```

```

RESULT 37
US-09-966-546-18
; Sequence 18, Application US/09966546
; Patent No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20020168716A1el Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; PRIORITY FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 815
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-546-18
```

```

Query Match      87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VRYED 5
        |||||
DB      376 VRYED 380
```

```

RESULT 38
US-09-966-545-18
; Sequence 18, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
```

;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 18
;; LENGTH: 815
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-545-18

Query Match 87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 376 VRYED 380

RESULT 39
US-09-965-212-18
;; Sequence 18, Application US/09965212
;; Publication No. US20030003462A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20030003462A1el Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/965,212
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: US/09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: USSN 60/128,514
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 18
;; LENGTH: 815
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-212-18

Query Match 87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 376 VRYED 380

RESULT 40
US-09-966-546-16
;; Sequence 16, Application US/0996546
;; Patent No. US20020168716A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20020168716A1el Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,546
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-546-16

Query Match 87.1%; Score 27; DB 9; Length 842;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 403 VRYED 407

RESULT 41
US-09-966-545-16
;; Sequence 16, Application US/09966545
;; Patent No. US20020172999A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,545
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-545-16

Query Match 87.1%; Score 27; DB 9; Length 842;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 403 VRYED 407

RESULT 42
US-09-965-212-16
;; Sequence 16, Application US/09965212
;; Publication No. US20030003462A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20030003462A1el Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/965,212
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: US/09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: USSN 60/128,514
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-212-16

Query Match 87.1%; Score 27; DB 9; Length 842;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
Db 403 VRYED 407

RESULT 43

US-09-808-602-58
; Sequence 58, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-58

Query Match 87.1%; Score 27; DB 9; Length 850;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
Db 406 VRYED 410

RESULT 44

US-09-808-602-73
; Sequence 73, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-73

Query Match 87.1%; Score 27; DB 9; Length 983;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VRYED 5
Db 539 VRYED 543

RESULT 45

US-10-013-136-2
; Sequence 2, Application US/10013136
; Publication No. US20030023067A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Elsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/10/013,136
; PRIOR FILING DATE: 2001-12-06
; CURRENT APPLICATION NUMBER: US/09/412,554
; PRIOR FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-136-2

Query Match 87.1%; Score 27; DB 9; Length 983;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
Db 539 VRYED 543

RESULT 46

US-09-782-714-2
; Sequence 2, Application US/09782714
; Patent No. US20020034808A1
; GENERAL INFORMATION:
; APPLICANT: Darst, Seth A
; APPLICANT: Zhang, Gongyi
; APPLICANT: Campbell, Elizabeth
; APPLICANT: Minakin, Leonid
; APPLICANT: Severinov, Konstantin
; TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
; FILE REFERENCE: 600-1-258
; CURRENT APPLICATION NUMBER: US/09/782,714
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 09/396,651
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Thermus aquaticus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (695)..(696)
; OTHER INFORMATION: Any amino acids can be at these two positions.
US-09-782-714-2

Query Match 87.1%; Score 27; DB 10; Length 1119;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5

Db 613 VRYED 617

```

RESULT 47
US-09-864-761-34609
; Sequence 34609, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/633,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34609
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007383.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BF474, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: SWISSPROT HIT: P28331, EVALUATE 9.00e-20

```

; OTHER INFORMATION: EST_HUMAN HIT: AW614813.1, EVALUATE 3.00e-19
US-09-864-761-34609

Query Match 83.9%; Score 26; DB 10; Length 45;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 26 VRYDDI 31

```

RESULT 48
US-09-764-878-127
; Sequence 127, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-878-127

```

Query Match 83.9%; Score 26; DB 10; Length 144;
Best Local Similarity 66.7%; Pred. No. 97;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 71 VRYDDI 76

```

RESULT 49
US-09-764-860-461
; Sequence 461, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 461
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-461

```

Query Match 83.9%; Score 26; DB 10; Length 144;
Best Local Similarity 66.7%; Pred. No. 97;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 1 VRYDDI 6

Db 71 VRYDDI 76

RESULT 50
 US-09-764-868-919
 ; Sequence 919, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 919
 ; LENGTH: 155
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-868-919

Query Match 83.9%; Score 26; DB 9; Length 155;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYDDL 6
 :||:|
 Db 89 IRYKDL 94

Search completed: February 20, 2003, 13:40:58
 Job time : 12.7429 secs

This Page Blank (uspto)